

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



INTERNATIONAL BUREAU OF PATENT COOPERATION
3, rue de la Harpe, CH-1015, Genève, Suisse
P.O. Box 68, CH-1015, Genève, Suisse
P.O. Box 68, CH-1015, Genève, Suisse

(43) International Publication Date
22 November 2001 (22.11.2001)

PCT

(10) International Publication Number
WO 01/88114 A2

(51) International Patent Classification: C12N 15/00

LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(21) International Application Number: PCT/EP01/05794

(22) International Filing Date: 18 May 2001 (18.05.2001)

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0012231.7 19 May 2000 (19.05.2000) GB

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a
patent (Rule 4.17(ii)) for the following designations AE, AG,
AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD,
GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,
MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW,
ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ,
UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU,
TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES,
FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI
patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR,
NE, SN, TD, TG)
— of inventorship (Rule 4.17(iv)) for US only

(71) Applicant (for all designated States except US): DEV-
GEN NV [BE/BE]; Technologiepark 9, B-9052 Zwij-
naarde (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PLAETINCK,
Geert [BE/BE]; Pontstraat 16, B-9820 Merelbeke (BE).
RENARD, Jean-Pierre [BE/BE]; Peter Benoîtlaan
141, B-9050 Gentbrugge (BE). MORTIER, Katherine
[BE/BE]; Lange Boomgaardstraat 25, B- Gent (BE).
BOGAERT, Thierry [BE/BE]; Wolvendreef 26g, B-8500
Kortrijk (BE).

(74) Agent: BOULT WADE TENNANT; Verulam Gardens,
70 Gray's Inn Road, London WC1X 8BT (GB).

Published:

— without international search report and to be republished
upon receipt of that report

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: GENE EXPRESSION SYSTEM

(57) Abstract: There are provided DNA constructs, including replicable cloning vectors and expression vectors, comprising a bac-
teriophage promoter operably linked to an outtron sequence. The expression vectors provided by the invention are useful in the
expression of recombinant polypeptides in host cells or organisms and are particularly useful in expression of recombinant polypep-
tides in nematode worms such as *C. elegans*.

WO 01/88114 A2

Best Available Copy

NUC 38884

- 1 -

GENE EXPRESSION SYSTEM

Field of the invention

5 The invention relates to the expression of DNA, genes, cDNAs, proteins, peptides and parts thereof in the nematode worm *C. elegans*. In particular, the invention relates to methods of improving the translation of RNAs transcribed in *C. elegans* using a
10 bacteriophage polymerase by introduction of a trans-splice recognition site recognised by an SL1 trans-splice recognition sequence into the DNA template transcribed by the bacteriophage polymerase.

15 Background to the invention

Eukaryotic versus prokaryotic expression.

 Bacteriophage RNA polymerases, such as T7, T3, and SP6, and their corresponding promoters have been
20 used extensively to drive the expression of heterologous genes in a variety of organisms. In co-pending International patent application No. WO 00/01846, Plaetinck et al. describe the use of the T7 system to express DNA, genes, cDNA, proteins and
25 peptides of parts thereof and for the expression of double-stranded RNA (dsRNA) in the nematode model system *C. elegans*.

 The bacteriophage expression systems are well known in the art for use in prokaryotic host cells,
30 such as *E. coli*, and have the advantage that they provide simple and strong expression systems dependent only on one RNA polymerase and one well defined promoter. The application of such efficient expression systems in eukaryotic organisms is,
35 however, not evident, mainly because messenger RNAs from eukaryotes and prokaryotes have a different

- 2 -

structure, which has implications for translation efficiency and RNA stability.

1 Messenger RNAs of higher eukaryotes share a functionally essential 5' CAP structure. This
5 structure is generated during a capping reaction that is linked exclusively to RNA polymerase II transcription. Prokaryotic RNA polymerases such as bacteriophage T3, T7 and SP6 polymerases do not
10 provide messenger RNAs with such a CAP structure, leading to inefficient translation in eukaryotic systems (Fuerst et al. J. Mol. Biol:206:333-348 (1989)).

One way to improve translation of uncapped mRNAs in eukaryotic systems is by the insertion of an
15 internal ribosome entry site (IRES) sequence 5' of the coding sequence. For example, Elroy-Stein, et al., Proc. Natl. Acad. Sci. USA 87:6743-6747 (1990), describe the cloning of the untranslated region of the ECMV virus downstream of the T7 promoter in order to
20 enhance the efficiency of translation. In other systems translation of T7-derived transcripts may be enhanced by addition of a CAP structure derived from a capped transcript. For example, in Trypanosoma a 5' CAP structure is added to T7 generated RNA transcripts
25 by a natural occurring trans-splicing reaction (Wirtz et al. NAR 22:3887-3894 (1994)).

Trans-splicing in *C. elegans*.

In *C. elegans* many mRNAs contain an identical
30 short leader sequence, designated the spliced leader (SL). This splice leader is donated by a small RNA (SL RNA) via a trans-splicing reaction. This trans splicing was first observed by Krause et al., Cell 49:753-61 (1987). The splice leader RNA exists as a
35 small nuclear ribonucleoprotein particle and has the trimethylguanosine cap that is characteristic of

- 3 -

eukaryotic small nuclear RNAs. The trimethylguanosine cap present on the spliced leader RNA is transferred to the pre-mRNA during the trans-splicing reaction. Thereafter, the trimethylguanosine cap is maintained
5 on the mature mRNA (Van Doren et al., Mol. Cell. Biol. 10:1769-1772 (1990). The trans-splicing signal for such a splice leader is essentially an intron missing only the 5' splice site, designated an 'outtron'. An outtron has essentially all the intron sequence
10 including a trans-splice acceptor site homologous to a UUUCAG sequence preceded by a AU rich region (Conrad et al., NAR 21:913-919 (1993). Introduction of an outtron into the 5' untranslated region of a *C. elegans* gene converts it to a trans-spliced gene (Conrad et
15 al., EMBO J. 12:1249-1255 (1993); Conrad et al. Mol. Cell Biol. 11:1931-1926 (1991)) and introduction of donor sites in a natural trans-spliced *C. elegans* gene prevents trans-splicing and converts it into a more conventional gene.

20

Description of the invention.

Until recently, expression of heterologous and homologous genes in *C. elegans* was mainly achieved by linking an appropriate coding sequence to a selected
25 *C. elegans* promoter. The present inventors have recently demonstrated that the recombinant gene expression in *C. elegans* can be based on the prokaryotic T7 expression system (WO 00/01846). However, the present inventors found that the
30 expression system was far from being efficient, or at least the resulting expression was much lower than would be expected from this T7 related expression system. It was concluded that this low expression was mainly due to RNA instability or translation arrest.
35 Furthermore, it was reasoned that fundamental differences between prokaryotic and eukaryotic

- 4 -

expression systems, particularly the requirement for capping of the 5' end of the mRNA for efficient translation in eukaryotic systems, was the main reason for this unexpectedly low expression.

5 The inventors have now developed a solution to the problem of the inefficiency of the T7 system in eukaryotic host cells and organisms, particularly in *C. elegans*, and have constructed a generally applicable expression system which allows for the
10 efficient expression of genes, DNA, cDNA, peptides and proteins under the regulation of the T7 promoter in *C. elegans*.

 Therefore, in accordance with a first aspect of the invention there is provided a DNA construct
15 comprising a bacteriophage promoter operably linked to an outtron sequence.

 It is an essential feature of the DNA construct of the invention that the bacteriophage promoter and the outtron sequence are "operably linked", that is to
20 say they are arranged in a relationship permitting them to function in their intended manner. In this case, the bacteriophage promoter is positioned upstream of the outtron sequence such that it is capable of promoting transcription of the outtron
25 sequence upon binding of an appropriate RNA polymerase, with the outtron sequence forming the extreme 5' end of the resulting transcript.

 The DNA construct may further comprise at least one restriction enzyme recognition site positioned
30 downstream of and proximal to the outtron sequence. Advantageously, the DNA construct may contain multiple restriction sites forming a multi-cloning site. The purpose of the restriction site/multi-cloning site is to facilitate cloning of a heterologous or homologous
35 DNA fragment downstream of the outtron sequence. A DNA construct comprising a bacteriophage promoter, an outtron sequence and a restriction site/multi-cloning

- 5 -

site may therefore be referred to hereinafter as an 'outtron cloning construct'.

! In an outtron cloning construct it is advantageous for the restriction site/multi-cloning site to be
5 positioned fairly proximal to the outtron sequence (e.g. within 100bp) such that a heterologous or homologous sequence inserted at this site may be co-transcribed with the outtron sequence on a single mRNA. However, further sequence elements may be interposed
10 between the outtron sequence and the restriction site/multi-cloning site. For example, the general purpose vector pDW3123 described in the accompanying examples has a synthetic intron A sequence between the outtron sequence and the multi-cloning site.

15 In one preferred embodiment of the invention, the DNA construct is a replicable cloning vector, such as, for example, a plasmid vector. In addition to the bacteriophage promoter, outtron sequence and optional restriction site/multi-cloning site, the vector may
20 further contain one or more of the general features commonly found in cloning vectors, for example an origin of replication to allow autonomous replication within a host cell and a selective marker, such as an antibiotic resistance gene. Although not essential,
25 the vector may also contain a poly-adenylation signal to stabilize and process the 3' end of the mRNA transcribed from the bacteriophage promoter. A preferred example is the 3'UTR from the *C. elegans* unc-54 gene, but any other 3'UTR or polyadenylation
30 signal may be used.

Outtron-containing DNA constructs according to the invention may be easily be constructed from the component sequence elements using standard recombinant techniques well known in the art and described, for
35 example, in F. M. Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994).

- 6 -

Outron sequences for use in the constructs of the invention may be isolated from natural *C. elegans* genes using standard molecular biology techniques. For example, a natural outtron sequence might be amplified using the polymerase chain reaction or an equivalent amplification technique using *C. elegans* genomic DNA as a template. Alternatively, synthetic outtron sequences may be synthesised, for example, by annealing two complementary single stranded oligonucleotides, as illustrated in the accompanying examples. Once a DNA fragment comprising the outtron sequence has been obtained it would be a matter of routine to assemble an outtron construct by linking the outtron in the correct orientation relative to the bacteriophage promoter.

The sequences of the commonly used bacteriophage promoters, e.g. T7, T3 and SP6, are well known in the art and oligonucleotides containing functional phage promoter sequences can be readily synthesised using standard oligonucleotide synthesis techniques. It would be a matter of routine to insert such a synthetic promoter sequence into, for example, a plasmid vector backbone containing, for example, an origin of replication a selective marker and a suitable restriction site. Alternatively, one of the many plasmid vectors containing bacteriophage promoter sequences known in the art may be used as the starting point for the construction of a plasmid-based outtron cloning vector. The known vectors generally contain, in addition to the phage promoter sequence, one or more restriction sites conveniently positioned downstream of the phage promoter and also a bacterial origin of replication and a selective marker. Once the vector backbone is in place the outtron sequence may simply be inserted in the appropriate position downstream of the bacteriophage promoter.

In a particularly useful embodiment the invention

provides a DNA construct for use in bacteriophage promoter-driven expression of a polypeptide in a eukaryotic host cell or organism. This construct comprises a bacteriophage promoter operably linked to
5 a DNA sequence such that it is capable of initiating transcription of the DNA sequence upon binding of an appropriate RNA polymerase to the promoter, wherein the aforesaid DNA sequence comprises an outtron sequence and at least one open reading frame
10 positioned downstream of the outtron sequence.

The open reading frame may be essentially any protein-encoding DNA sequence bounded by start and stop codons. This protein-encoding DNA sequence may include introns, as both trans-splicing and cis-
15 splicing can occur together.

A DNA construct according to this embodiment of the invention, which may be referred to hereinafter as an 'outtron expression construct', may be derived from an outtron cloning construct by insertion of a
20 heterologous or homologous protein-encoding DNA fragment into the restriction site/multi-cloning site. It is essential that the heterologous or homologous DNA fragment be inserted downstream of the outtron sequence such that the two sequences may be co-
25 transcribed, with the outtron sequence forming part of the 5' untranslated region of the resulting mRNA.

The outtron expression construct may advantageously form an expression vector, such as, for example, a plasmid vector. Most preferably, the
30 expression vector will be one suitable for use in the nematode worm *C. elegans*. In addition to the bacteriophage promoter, outtron sequence and protein-encoding DNA sequence (open reading frame), the expression vector may further contain one or more of
35 the general features commonly found in expression vectors, for example an origin of replication to allow autonomous replication within a bacterial host cell

and a selective marker, such as an antibiotic resistance gene. The vector may also contain a poly-adenylation signal to stabilize and process the 3' end of the mRNA transcribed from the bacteriophage promoter. A preferred example is the 3'UTR from the *C. elegans* unc-54 gene, but any other 3'UTR or polyadenylation signal may be used. An additional element, such as for example a synthetic intron, may be interposed between the outtron sequence and the open reading frame.

It is important that the open reading frame is positioned downstream of and proximal to the outtron sequence in the expression construct such that (i) the two elements are co-transcribed to form a single mRNA and (ii) the outtron sequence forms part of the 5' untranslated region of the mRNA. If the appropriate splicing machinery and a supply of SL RNAs is provided by the eukaryotic host cell or organism then the uncapped 5' end of the pre-mRNA transcribed from the expression construct will be replaced with a capped splice leader via the trans-splicing reaction. This will greatly increase the efficiency of translation in a eukaryotic host system.

The use of an outtron sequence at the extreme 5' end of the RNA provides a solution to the problem of reduced expression efficiency in eukaryotic systems wherever the type of promoter/polymerase used to drive gene expression leads to the production of uncapped transcripts, provided that the host cell or organism produces the spliced leader RNAs required for the trans-splicing reaction.

Outtron sequences which may be utilised in accordance with the invention include naturally occurring outtron sequences isolated from SL1-specific *C. elegans* genes (Conrad, R. Functional analysis of a *C. elegans* trans-splice acceptor. *Nucleic Acids Res.* 1993, 21(4), pp913-919; Conrad, R. SL1 trans-splicing

- 9 -

specified by AU-rich synthetic RNA inserted at the 5' end of *Caenorhabditis elegans* pre-mRNA. *RNA*. 1995, 1(2), pp164-170) and also synthetic outtron sequences which are functionally equivalent to the natural *C. elegans* outtron sequences, including variants of naturally occurring *C. elegans* outtrons. The phrase "functionally equivalent" means that the synthetic intron is recognised by the *C. elegans* trans-splicing machinery and can be trans-spliced to a *C. elegans* splice leader RNA, preferably the SL1 splice leader.

Experimental evidence indicates that trans-splicing in *C. elegans* is signalled by an AU-rich intron-like sequence followed by a splice acceptor site (Conrad *et al* 1993 and 1995). For the purposes of the present application the terms "outtron" or "outtron sequence" should be interpreted as referring to both the AU-rich region from the 5' end of the pre-mRNA to the trans-splice acceptor site and the trans-splice acceptor site itself. In connection with the DNA constructs of the invention, the terms "outtron" and "outtron sequence" refer to features present in the DNA which encodes the pre-mRNA.

The consensus splice acceptor site for trans-splicing of outtrons and the consensus 3' splice acceptor site for cis-splicing of introns are essentially identical (UUUCAG). Moreover, a normally trans-spliced acceptor site can be efficiently cis-spliced when a donor splice site is inserted upstream within the outtron sequence. It is therefore important that the outtron constructs described herein do not contain any potential splice donor sequence upstream of the splice acceptor within the outtron and downstream of the transcription start site such that it will be transcribed in the mRNA encoded by the construct. If such a site were present then there would be a potential for cis-splicing rather than

- 10 -

trans-splicing.

It has also been observed that the overall length of the outtron has an effect on the efficiency of trans-splicing, longer outtrons in general working better than shorter ones (Conrad *et al.* 1995). Advantageously, the outtron sequences for inclusion into the outtron constructs described herein should be greater than about 50nt in length.

A synthetic outtron containing an AT stretch and a TTTTCAG sequence has been shown to be functional in *C. elegans*. As illustrated in the accompanying Examples, the insertion of an outtron sequence into the 5' untranslated region of GFP reporter construct, downstream of the promoter and upstream of the GFP open reading frame, is required for optimal expression of bacteriophage RNA polymerase transcribed reporter gene mRNA in *C. elegans*.

Suitable bacteriophage promoters which may be used in the DNA constructs according to the invention include T7, T3 and SP6 promoters, with T7 being the most preferred. As discussed above, these bacteriophage promoters have long been known to be useful tools in molecular biology since they can provide simple and strong expression systems dependent only on the binding of the specific or cognate RNA polymerase.

In a still further aspect, the invention provides a method for expressing a recombinant polypeptide in *C. elegans*, which method comprises:

introducing an outtron expression construct, as described above, said construct being an expression vector suitable for use in *C. elegans*, into a *C. elegans* strain which expresses an RNA polymerase specific for the bacteriophage promoter present in said DNA construct in one or more tissues or cell

- 11 -

types.

An outtron expression vector for use in this method may be constructed by inserting DNA encoding the polypeptide of interest into an outtron cloning
5 vector, as described above. The vector must be one which is suitable for use in *C. elegans*, plasmid-based vectors are the most preferred.

The *C. elegans* worms are preferably transgenic worms carrying a transgene capable of expressing the
10 RNA polymerase in one or more tissues or cell types. The term "transgene capable of expressing" as used herein means a nucleic acid molecule comprising a nucleotide sequence encoding the polymerase operably linked to a promoter. The promoter may be any
15 promoter which functions in *C. elegans* and may be general (i.e. active in substantially all tissues and cell types), tissue-specific, cell type-specific, constitutive, inducible etc. Most preferably, the promoter will exhibit tissue or cell type-specificity.
20 With the use of a tissue or cell type-specific promoter of the appropriate specificity it is possible to control the site of RNA polymerase expression within *C. elegans* and hence control the site of expression of the recombinant polypeptide.

25 Methods for the construction of transgenic *C. elegans* worms are known in the art and are particularly described by Craig Mello and Andrew Fire, *Methods in Cell Biology*, Vol 48, Ed. H.F. Epstein and D.C. Shakes, Academic Press, pages 452-480.

30 In a further aspect the invention provides a kit for use in recombinant expression of a polypeptide in *C. elegans*, the kit comprising an outtron cloning construct, as described above, and optionally a supply
35 of *C. elegans* nematode worms expressing an RNA polymerase specific for the bacteriophage promoter

- 12 -

present in the said outtron cloning construct in one or more tissues or cell types.

The kit might further contain control inserts and control constructs, e.g. a reporter gene inserts and constructs which could be used to check efficiency of cloning steps and transfection steps, respectively. It might also contain constructs which may be used as selectable markers in the transfection procedure, e.g. a rol 6 plasmid (see below).

The invention further provides methods for the construction of transgenic *C. elegans* expressing a recombinant polypeptide in one or more tissues or cell types. One such method comprises introducing an outtron expression construct, as described above, said construct being an expression vector suitable for use in *C. elegans* comprising an open reading frame encoding the desired recombinant polypeptide, into a *C. elegans* strain which expresses an RNA polymerase specific for the bacteriophage promoter present in said DNA construct in one or more tissues or cell types, and isolating transgenic *C. elegans* lines which stably express the said polypeptide. The *C. elegans* strain expressing the polymerase is preferably a transgenic strain carrying a transgene capable of expressing the RNA polymerase in one or more tissues or cell types, as described above. As aforesaid, transgenic *C. elegans* lines can readily be constructed using standard techniques well known in the art.

In an alternative approach, the method may comprise introducing into a background *C. elegans* strain (i) an outtron expression construct, as described above, said construct being an expression vector suitable for use in *C. elegans* comprising an open frame encoding the desired recombinant polypeptide, and (ii) a DNA construct suitable for expression of an RNA polymerase specific for the

- 13 -

bacteriophage promoter present in the outtron expression construct in one or more tissues or cell types of *C. elegans*, and isolating transgenic *C. elegans* lines which stably express the said polypeptide. The second DNA construct may, advantageously, be an expression vector comprising a nucleotide sequence encoding the polymerase operably linked to a promoter having the appropriate tissue or cell type specificity.

10 In carrying out the methods of the invention one may employ standard techniques well known in the art for construction and selection of transgenic *C. elegans* lines. Such techniques are described, for example, in techniques described in Methods in Cell Biology, vol 84; *Caenorhabditis elegans*: modern biological analysis of an organism, ed. Epstein and Shakes, academic press, 1995. Foreign DNA (e.g. plasmid DNA) may be introduced into *C. elegans* using microinjection or ballistic transformation, as described in the applicant's co-pending International patent application No. WO 99/49066. In order to facilitate the selection of transgenic strains a marker plasmid may be co-introduced with the transgenes. A typical example is the plasmid pRF4 (Mello, C. C. et al. EMBO J. 10, 3959-3970 (1991)) which carries the rol-6 gene. *C. elegans* expressing rol-6 can be identified by screening for the roller phenotype. Any other *C. elegans* dominant selectable phenotypic marker, of which there are many known in the art, may be used to facilitate selection of transgenic lines. A useful example is green fluorescent protein (or any of the equivalent autonomous fluorescent proteins known in the art).

30 In a still further aspect the invention provides transgenic *C. elegans* worms which contain an outtron expression construct, as described above, said

- 14 -

construct being an expression vector suitable for use in *C. elegans*, and which further express an RNA polymerase specific for the bacteriophage promoter present in the outtron expression construct in one or more tissues or cell types.

The present invention will be further understood with reference to the following non-limiting Examples, together with the accompanying drawings in which:

10

Figure 1 illustrates the construction of a T7-outtron-GFP vector. (A) sequence of the synthetic outtron produced by annealing oligonucleotides o-GN59 and o-GN60. (B) summary of the strategy used to construct vector pDW3124.

15

Figure 2 shows plasmid maps for pDW3123 (outtron cloning vector) and pDW3124 (outtron expression vector for GFP expression).

20

Figure 3 is a plasmid map of pGN148 which contains a T7 RNA polymerase coding sequence under the regulation of the *C. elegans* SERCA promoter.

25

Figure 4 illustrates the nucleotide sequence of pGN148.

30

Figure 5 illustrates the nucleotide sequence of pDW3123 annotated to show the positions of the T7 promoter, outtron, synthetic intron A, multi-cloning site and unc-54 3' UTR sequences and also the ampicillin resistance gene.

35

Figure 6 illustrates the nucleotide sequence of pDW3124 annotated to show the positions of the T7 promoter, outtron, synthetic intron A, GFP with introns and unc-54 3' UTR sequences and also the ampicillin

- 15 -

resistance gene.

Example 1 -Construction of a T7-outtron-GFP containing vector (pDW3124)

5 A SL1 trans-splice acceptor site (outtron) was cloned into a vector downstream of the T7 promoter and upstream of the GFP to be expressed.

 A synthetic outtron consisting of two partially overlapping oligonucleotides (o-GN59 and o-GN60, see
10 Figure 1) was inserted into a XbaI/XmaI digested T7 promoter GFP construct. Briefly, 25µl o-GN59 and 25µl o-GN60 (100µM) were denatured for 5 minutes at 94°C, annealed for 30 minutes at 68°C then cooled to 4°C. 1µl of XmaI/XbaI digested pDW3120 and 10µl of the
15 annealed oligos were then ligated using T4 ligase overnight at 16°C, transformed into competent *E. coli* and analysed by restriction digestion and DNA sequencing, all according to standard molecular biology procedures. The resulting vector was
20 designated pDW3124 (Figures 1 and 2).

 The outtron contains an AU rich sequence followed by a splice-acceptor site as described by Conrad et al, NAR 21:913-919 (1993) (see Figure 1).

25 Example 2-Construction of a T7-Outtron MCS vector

 A general purpose vector was constructed to facilitate expression of other DNA sequences in *C. elegans* under the control of the T7 promoter. This was done by digesting vector pDW3124 with HindII
30 (position 179) and PvuII (position 1029) (partial digest) and re-ligating the blunt ends, resulting in vector pDW3123 (Figure 2).

35

- 16 -

Example 3-The expression of heterologous genes in *C. elegans* regulated by the T7 promoter requires trans-splicing.

5 Wild-type *C. elegans* nematodes were co-injected with various combinations of the following test plasmids:

- 1) GFP reporter plasmid
GFP: pDW2020
10 outtron-GFP: pDW2024
T7 promoter-GFP: pDW3120
T7 promoter-outtron-GFP: pDW3124
- 2) T7 polymerase expression plasmid SERCA T7
15 polymerase: pGN148 together with pRF-4 (rol-6) as marker.

For every co-injection experiment, a total concentration of 200 ng DNA/ μ l was used (plasmid concentration was 50 ng/ μ l and carrier DNA was added up to 200ng/ μ l). For every co-injection \pm 15 adult worms were injected.

25 F1 offspring showing the marker rol-6 phenotype were isolated and then selected for further study. The next generation (F2) of the roller lines were screened for GFP expression in the pharynx, vulva, tail and body wall muscles. These are the tissues in which the bacteriophage T7 RNA polymerase is known to
30 be expressed when under the control of the *C. elegans* SERCA promoter (as in the construct pGN148)

The results are shown in Table 1 below, which indicates the number of lines expressing GFP vs total number of lines isolated.

35

- 17 -

	1	2	3
A	Construct	no T7-polymerase construct	with T7-polymerase construct (50ng) pGN148
B	GFP (50ng) pDW2020	0/8	2/6*
C	outtron::GFP (50ng) pDW2024	0/11	3/8*
5 D	T7-promoter::GFP (50ng) pDW3120	0/3	0/5
E	T7-promoter::outtron::GFP (50ng) pDW3124	0/7	13/13

10 * GFP-expression most probably result of recombination
in the extrachromosomal array

15 No GFP expression was observed in the experiments
where the T7 RNA polymerase was absent (cells B2, C2,
D2, E2).

In the experiments where the T7 RNA polymerase
expressing vector was co-injected with GFP vectors
without a T7 promoter, as in the cells B3 and C3, GFP
expression was sometimes observed. This is probably
20 due to recombination events in the extrachromosomal
arrays, resulting in transcription of GFP directly
from the SERCA promoter.

25 In the experiments where the T7 promoter-GFP construct
and the SERCA T7 RNA polymerase were co-injected, no
GFP expression could be observed (cell D3). In
contrast, all of the lines isolated from the
experiments where the GFP transcript contained an
outtron at its 5' site (n=13) expressed GFP (cell E3).
30 The outtron is a favourable target for SL1
trans-splicing. Since SL1 RNA molecules contain a 5'

- 18 -

trimethylguanosine CAP structure which is transferred to the mature mRNA this results in improved translation of the RNA and hence better expression of GFP. Without the outtron the T7 RNA polymerase transcripts do not carry a CAP structure at their 5' end, leading to inefficient translation. The results of this experiment illustrate the importance of trans-splicing for efficient expression of heterologous and homologous genes transcribed by prokaryotic polymerases in *C. elegans*.

SEQUENCE LISTING

SEQ ID NO: 1 Oligonucleotide o-GN59
15 SEQ ID NO: 2 Oligonucleotide 0-GN60
SEQ ID NO: 3 Plasmid pDW3123
SEQ ID NO: 4 Plasmid pDW3124
SEQ ID NO: 5 Plasmid pGN148

- 19 -

Claims:

1. A DNA construct comprising a bacteriophage promoter operably linked to an outtron sequence.

2. A DNA construct as claimed in claim 1 which further comprises at least one restriction enzyme recognition site positioned downstream of and proximal to the outtron sequence.

3. A DNA construct as claimed in claim 2 which comprises a multi-cloning site positioned downstream of and proximal to the outtron sequence.

4. A DNA construct as claimed in claim 2 or claim 3 which further comprises a DNA fragment inserted at the said restriction site or at a restriction site within the said multi-cloning site.

5. A DNA construct as claimed in any one of claims 1 to 4 which is a replicable cloning vector.

6. A DNA construct as claimed in any one of claims 1 to 5 wherein the outtron sequence comprises a 3' splice acceptor site having the sequence TTTCAG preceded by an AT-rich region.

7. A DNA construct as claimed in claim 6 wherein the outtron sequence comprises the nucleotide sequence illustrated in Figure 1A.

8. A DNA construct as claimed in any one of claims 1 to 7 wherein the bacteriophage promoter is the T7, T3 or SP6 promoter.

9. A DNA construct for use in bacteriophage promoter-driven expression of a polypeptide in a

- 20 -

eukaryotic host cell or organism, which construct comprises a bacteriophage promoter operably linked to a DNA sequence such that it is capable of initiating transcription of said DNA sequence upon binding of the appropriate RNA polymerase to the promoter, wherein the said DNA sequence comprises an outtron sequence and at least one open reading frame positioned downstream of the outtron sequence.

10 10. A DNA construct as claimed in claim 9 which is an expression vector.

11. A DNA construct as claimed in claim 9 or claim 10 wherein the outtron sequence comprises a 3' splice acceptor site having the sequence TTTCAG preceded by an AT-rich region.

12. A DNA construct as claimed in claim 11 wherein the outtron sequence comprises the nucleotide sequence illustrated in Figure 1A.

13. A DNA construct as claimed in any one of claims 9 to 12 wherein the bacteriophage promoter is the T7, T3 or SP6 promoter.

14. A kit for use in recombinant expression of a polypeptide in *C. elegans*, the kit comprising a DNA construct as claimed in any one of claims 1 to 3, and optionally *C. elegans* worms expressing an RNA polymerase specific for the bacteriophage promoter present in said DNA construct in one or more tissues or cell types.

15. A method for expressing a recombinant polypeptide in *C. elegans* which method comprises: introducing a DNA construct as claimed in any one

- 21 -

of claims 9 to 13, said construct being an expression vector suitable for use in *C. elegans*, into a *C. elegans* strain which expresses an RNA polymerase specific for the bacteriophage promoter present in
5 said DNA construct in one or more tissues or cell types.

16. A method of generating transgenic *C. elegans* expressing a recombinant polypeptide, which method
10 comprises:

introducing a DNA construct as claimed in any one of claims 9 to 13 comprising an open reading frame encoding the recombinant polypeptide, said construct being an expression vector suitable for use in *C. elegans*, into a *C. elegans* strain which expresses an
15 RNA polymerase specific for the bacteriophage promoter present in said DNA construct in one or more tissues or cell types, and

isolating transgenic *C. elegans* lines which
20 stably express the said polypeptide.

17. A method of generating transgenic *C. elegans* expressing a recombinant polypeptide, which method comprises:

25 introducing into *C. elegans* (i) a first DNA construct as claimed in any one of claims 9 to 13 comprising an open reading frame encoding the recombinant polypeptide, said construct being an expression vector suitable for use in *C. elegans*, and
30 (ii) a second DNA construct suitable for expression of an RNA polymerase specific for the bacteriophage promoter present in the first DNA construct in one or more tissues or cell types of *C. elegans*, and

isolating transgenic *C. elegans* lines which
35 stably express the said polypeptide

- 22 -

18. Transgenic *C. elegans* which contain a DNA construct as claimed in any one of claims 9 to 13, said construct being an expression vector suitable for use in *C. elegans*, and which further express an RNA
5 polymerase specific for the bacteriophage promoter present in said DNA construct in one or more tissues or cell types.

FIG. 1(A).

XbaI overhang	SspI	3' splice acceptor	
CTAGATTACAAC TAATTATACTTATTTGAATATTCAAATTTTCAGAC			o-GN59
TAATGTTGATTAATATGAATAAACTTATAAGTTTAAAAGTCTGGGCC			o-GN60
		XmaI overhang	

FIG. 1(B).

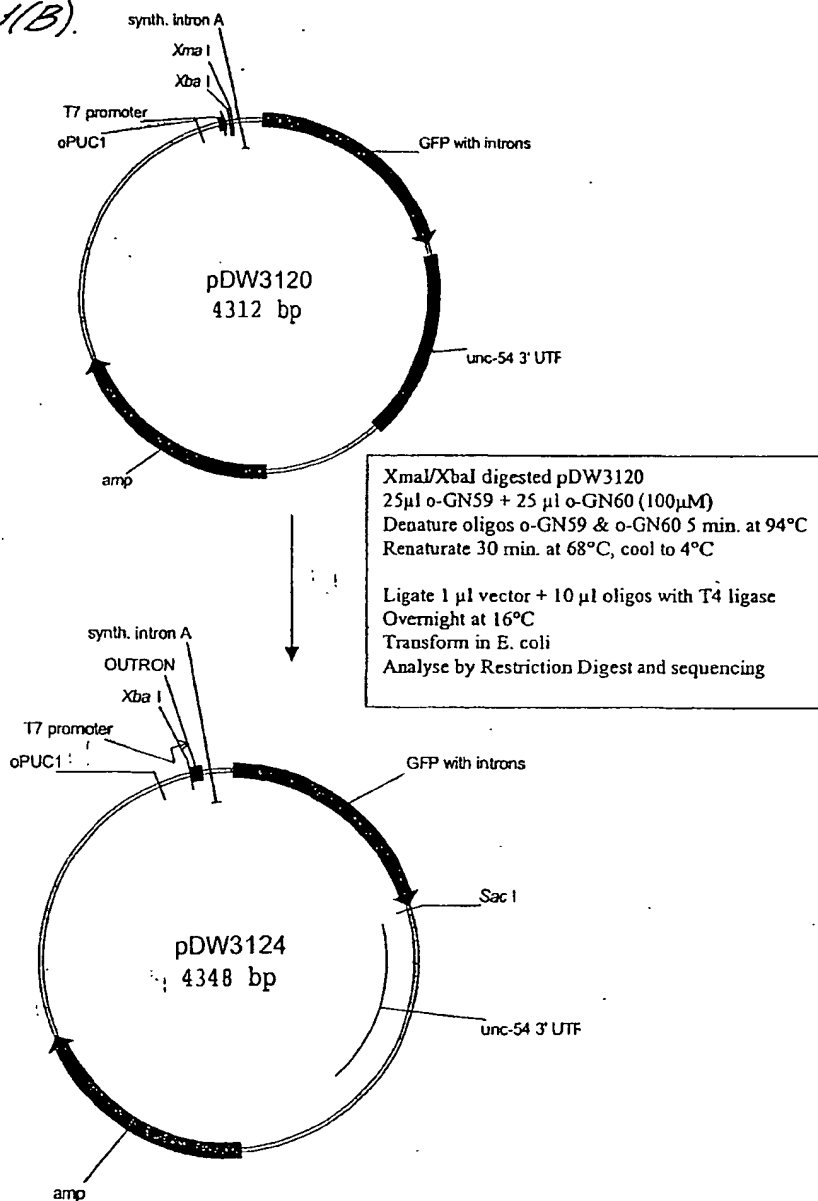


FIG. 2.

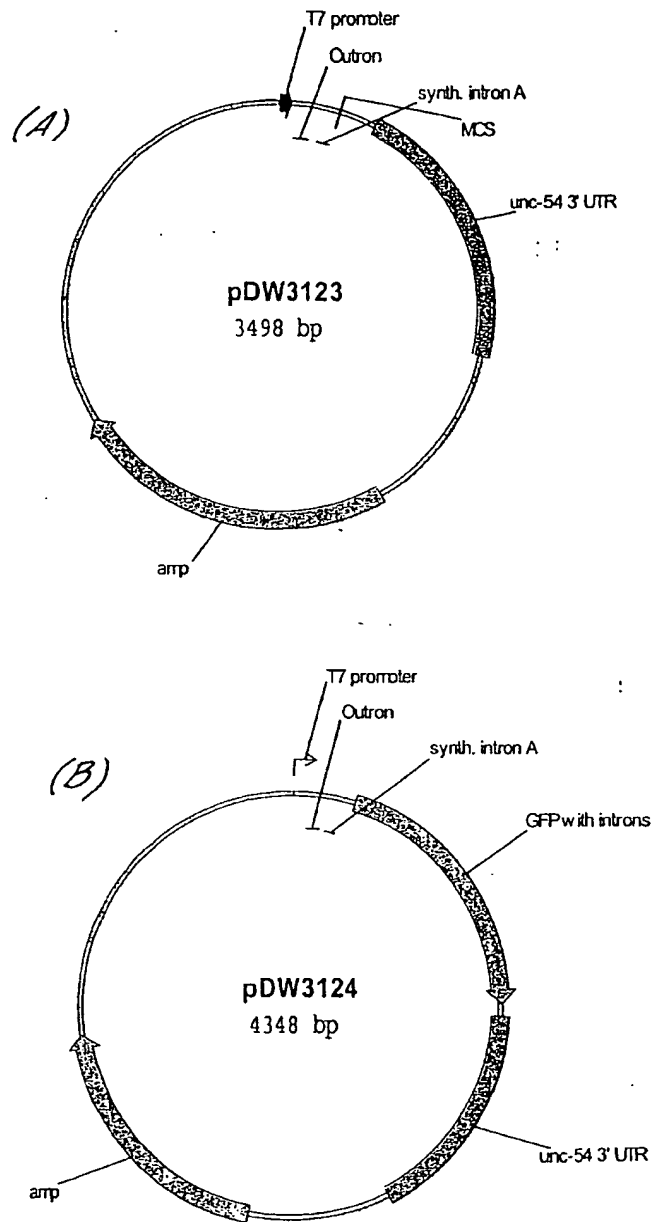


FIG. 3.

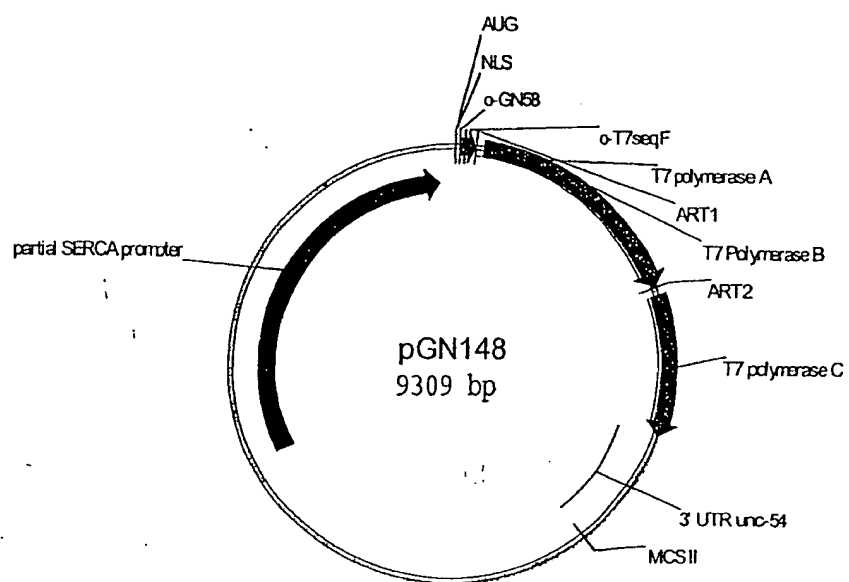


FIG. 4.

Nucleotide sequence of pGN148

atgactgctccaaagaagaagcgttaagggtaccggttaatgaacacgattaacatcgctaagaacgacttctc
tgacatcgactggctgctatccccgttcaacactctggctgaccattacgggtgagcgtttagctcggtaag
tttaaacatctagataactaactaacgattaacatttaaattttcagcgaacagttggcccttgagcatgag
tcttacgagatgggtgaagcacgcttccgcaagatggttgagcgtcaacttaagcgtggtaggttgcgga
taacgctgccgccaagcctctcatcactaccctactccctaagatgattgcacgcatcaacgactgggttg
aggaagtgaagcctaagcgcggcaagcgcggcagacgcttccagttcctgcaagaaatcaagccggaagcc
gtagcgtacatcaccatlaagaccactctgggttgccctaacagtgctgacaatacaaccggttcaggctgt
agcaagcgaatcggtcgggccattgaggacgaggctcgcttcgggtcgatccgtgacctgaagcgaagc
acttcaagaaaaacgttgaggaacaactcaacaagcgcgtagggcacgctctacaagaaagcatttatgcaa
gttgtcgaggctgacatgctcttaagggtctactcgggtggcaggcgtgggtctcgtggcataaggaaga
ctctattcctgtaggagtagcgtgcatcgagatgctcatlgagtcaaccggagtggttagcttacaccgcc
aaaatgctggcgttaggtcaagactctgagactatcgaactcgacacctgaatacgtgaggctatcgca
acccgtgcagggtgctgctggcgtatctctccgatggttccaaccttcgctagtttccctcctaaagcgtggac
tggcattactgggtgggtgctatgggctaacgggtcgctcgctcctggcgctgggtgctgactcacagtaaga
aagcactgctgcgtacgaagacgtttacatgctgagggtgtacaaagcgatlaacattgcgcaaaacacc
gcatggaaaaatcaacaagaaagtcctagcggctgcgaacgtaataaccaagtggaagcattgtccggctcga
ggacatccccgtcgattgagcgtgaagaactccccgatgaacccggaagacatcgacatgaatcctgaggctc
tcaccgcgtggaaacgtgctgcccgtgctgtgtaccgcaaggacaaggctcgcaagttctcgccgtatcagc
cttgagttcattgcttgagcaagccaaataagtttgctaacctaaggccatctgggtcccttacaacatgga
ctggcgcggtcggtgttacgctgtgtcaatgttcaaccgcgaagctaacgatatgacaaaaggactgctta
cgctggcgaaaggtaaaccaatcggttaaggaaggttactactgggtgaaaaatccacgggtgcaaacgtgtgcg
gggtgcgataaaggtcccgttcccctgagcgcacatcaagttcattgaggaaaaaccacgagaacatcatggcttg
cgctaagttctccactggagaacacttgggtgggtgagcaagattctccgttctgcttcccttgcttctgct
ttgagtacgctgggttacagcaccacggcctgagctataactgctcccttccgctggcgtttgacgggtct
tgctctggcatccagcacttctccgctgctcctcgagatgaggttaggtgggtcgcgcggttgtaagtttaaa
ctctatcctactaactaacgaagccttatttaaattttcagaacttgcttccctagtgaaaccggttcaggaca
tctacgggattgttgctaagaaagtcacagagattctacaagcagacgcaatcaatgggaccgataaacgaa
gtagttaccgtgaccgatgagaacactgggtgaaatctctgagaaggtcaagctgggcactaaggcactggc
tgggtcaatggcgtggcttacgggtgttactcgcagtgtagtaagcgttcagtcagcgtgggttacgggt
ccaaagagttcggctccgtcaacaagtgctggaagataccattcagccagctattgattccggcaagggt
ctgatgttactcagccgaatcagggtgctggatacatggctaagcgtgatttgggaatctgtgagcgtgac
gggtgtagctgggttgaaagcaatgaactgggttaagttcgtgctaaagctgctgggtgctgaggtcaaa
ataagaagactggagagattcttcgcaagcgttgcgctgtgcatgggtcactccggatgggttccctgtg
tggcagggaatacaagaagcctattcaaacgcgtttgaacctgatgttccctcggtcagttccgcttacagcc
taccattaaacaccaacaaagatagcgagattgatgcacacaaacaggagctgggtatcgctcctaaccttg
tacacagccaagacggtagccaccttcgtaagactgtagtggtgggcacacgagaagtaggaatcgaatct
tttgactgatttcagcactccttcgggtaccattccggctgacgctgcgaacctgttcaaacgagtgccgca
aactatgggtgacacatatgagtccttgtagtgtagtggtgatttctacgaccagttcgctgaccagttgc
acgagttcctaattggacaaaatgccagcacttccggctaaaggtaacttgaacctccgtgacatcttagag
tcggacttcgcgttcgcgtagaatccaactgagcgcgggtcgctaccattaccaacttgctcgtggtca
aaaataataggggcgctgctcatcagagtaggtttaaactgagttctactaactaacgagtaataatttaaa
ttttcagcatctcgcccccgtgctctgacttctaagtccaattactcttcaacatccctacatgctcttt
ctccctgtgctcccccctctatttttgttattatcaaaaaaacttcttcttaattctctgttttttagc
ttcttttaagtcacctctaacaatgaaattgtgtagattcaaaaaatagaatttaattcgtaataaaaagtcg
aaaaaaattgtgctccctccccctattataataattctatcccaaaatctacacaattgttctgtgtacac
ttcttatgttttttttacttctgataaatttttttgaacatcatagaaaaaacgcacacaaaataacct
tatcatatgttacgtttcagtttatgaccgcaatttttattcttcgcacgctctgggctctctatgacgtc
aaatcatgctcatcgtgaaaaagttttggagatttttgggaatttttcaatcaagtgaaagtttatgaaat
taattttctgcttttcttttttgggggtttccctattgtttgtcaagagtttcgaggacggcggttttc
ttgctaaaatcacaagatttgatgagcagatgcaagaaagatcggaagaaggtttgggtttgagggtcag
tgggaaggtgagtagaagttgatatttgaagtgagtagtgcttatgggggttttgccttaaatgacaga
atacatcccaatataccaaacataactgtttcctactagtcggccgtacgggccccttctgctcgcgcgt
ttcgggtgatcgggtgaaaaacctctgcacatgcagctcccgagacgggtcacagcttgctctgtaagcgga
tgccgggagcagacaagccgtcagggcgctcagcgggtgttgccgggtgtcgggggtgggttaactatg
cgcatcagagcagattgtactgagagtgacacatatgcggtgtgaaataccgcacagatgctgaagga

FIG. 4 (CONTINUED 1.)

aaataaccgcacagcgccgcttaaggccctcgtgatacgcctatTTTTataggttaatgcatgataataa
tggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttggtttatTTTctaa
atacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaa
gagtatgagattcaacatttcggtgcgccttatccctttttgcggcattttgccttcctgttttg
ctcaccagaaacgctgggtgaaagttaaagatgctgaagatcagttgggtgcacgagtggttacatcgaa
ctggatctcaacagcggttaagatccttgagagttttcgccccgaagaacgtttccaatgatgagcacttt
taaagttctgtatgtggcgcggtattatcccgatttgacgccccgaagagcaactcggtcgccccatc
actattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagta
agagaattatgcagtgctgcatataccatgagtgataaactgcggccaacttactctgacaacgactcgg
aggaccgaaggagctaacgcctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaac
cgagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgctg
cgcaaacctataactggcgaaacttactctagcttcccggaacaataatagactggatggaggcgga
taaagtgcaggaccacttctgcgctcggcccttcgggtgggtttattgctgataaaatcggagcg
gtgagcgtgggtctcgcggtatcatgcagcactggggccagatggtaagccctcccgatcgtagtattc
tacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgat
taagcatttggttaactgtcagaccagtttactcataataactttagattgatttaaaacttcattttaat
ttaaaggatctaggtgaagatcctttttgataatctcatgacaaaaatcccttaacgtgagtttcgttc
cactgagcgtcagacccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcgcgtaatctg
ctgcttgcaaaaaaaaccaccgctaccagcggtgggttggttgccggatcaagagctaccaactcctt
ttccgaaggtaactggcttcagcagagcgagataccaaatactgctcctctagtgtagccgtagttaggc
caccacttcaagaactctgtgacccgctacatacctcgtctgctaatcctgttaccagtggtgctgc
cagtgccgataagtcgtgtcttaccgggttggaactcaagacgatagttaccggataaggcgagcggtcgg
gctgaacggggggttcgtgcacacagcccagcttgaggcgaacgacctacaccgaactgagatacctacag
gtgagcattgagaagcgccagcttccgaagggaagggcggaacaggtatccggtaagcgcgaggt
cggaacaggagagcgacaggggagcttccagggggaacgcctgggtatctttatagtcctgctgggttc
gccacctctgacttgagcgtcgatttttgatgctcgtcagggggcgaggccctatggaaaaacgccagc
aacgcgcccttttacgggttcctggccttttgctggccttttgctcacatgttcttccctgcgttatcccc
tgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgcgcagccgaacgcagagc
gcagcagtgactgagcgaaggaagcggaagagcgcccaatacgcgaacccgctctccccgcgcttggcgcg
attcattaatgcagctggcagcagaggttcccgactggaaagcgggcagtgagcgcaacgcaattaatgt
gagttagctcactcattagccacccagcctttacactttatgcttccggctcgatggttggtggaattg
tgagcggataacaatttcacacaggaacagctatgacctgattacgccaagctgtaagtttaaacatga
tcttactaactaactattctcatttaaattttcagagcttaaaaatggctgaaatcactcacaacgatgga
tacgctaacacttggaatgaaataagcttgcatgctgcagagcaaaaaaatactgcttttctctgca
aatctgggtgctttctcaagagaaacttttgaaagtcggcgagcatttctcttctgacttctctctt
ccgcaaaaagcctagcatttttattgataatttgattacacacactcagagttcttgacatgataaagt
gtttcattggcactcgccttaacagttacatgacaagggcggtattattatcgatcgatattgaagcaaac
ccaaatgtgtgctcattttggagccccgttggggcagctgctctcaatataactagggagacgagga
ggggaccttatcgaaagtcgcatgagccattcttctcttttatgactctcttctcactctctcacacatta
atcgattcatagactcccatattccttgatgaaggtgtgggtttttagcttttttcccgatttgtaaaag
gaagaggtgacgatgttaggaaaaagagaacggagccgaaaaaacatccgtagtaagcttctctttaaag
ccgacactttttagacagcatctgcgcgtagttttgaagtttaaattttaaaaaataaaaaattagtttcaa
tttttttaattactaaaataggcaaaagtttttcaagaactctagaaaaactagcttaattcatgggtac
tagaaaaattctgttttaaatattatcttaagatgtaattacgagaagctttttgaaaaattct
caattaaaagaatttgccgattagaataaaaagcttccagaatgagtaaaaagctcaaattagaagttgt
ttttaaaggaaaaacacgaaaaaagaacactatttatcttttctccccgcgtaaaaatagttgtgtgat
aatagtgatccgctgtctatttgcaactcggtcttccacaccgtcttctctcacttgacccaacaggaaa
aaaaaacatcagctctgagacgggtgaattgccttatcaagagcgtcgtctcttcccccagtaacaaaaaa
aatttgggttcttactttatatttgtaggtcacaaaaaaaagtgatgacgttttggtgggtcgggtgt
ctccacaccctccgctccagcagcacacaatcatctcgtgtgtctcgcagcatccttgtagtgcgcg
ggctgtagatgaccacatcgcagcgcgaactacacaccacactcactttcgggtggtattactacacgtca
tcgttgtctgtagtctccgctcttctcgtccccactcactctcattattcccttgggtgtattgattttt
tttaaatggtagaccactcctgacgtttctacctttctgttttccgctccatttagattttatctggaat
tttttaaattttaggcagagagttctagttctgtttcttaaaagtctagtcagacatacatttctct
tctcatcaaaaaaaggttgataaaagaaactgggtatttcagaaagagtgctcgtctgtaaatgattca
aaaaaaattccccccctcgttgttttctcaaaatagagatcaacgggaatttttctctctcgattcaat
ttttgctgcgctctgctgctccaaagtgtgtgtgctccgagcaaaagatgagagaatttacaacagaaatg

NUC 38912

FIG. 5.

```

=====
T7 promoter                                     Outron
=====
1  AGCTTGGCGC CTAATACGAC TCACTATAGG GCTGCAGGTC GACTCTAGAT TACAACAAAT TATACTTATT
   TCGAACCGCG GATTATGCTG AGTGATATCC CGACGTCCAG CTGAGATCTA ATGTTGATTA ATATGAATAA

Outron                                     synth. intron A
=====
71  TGAATATTCA AATTTTCAGA CCCGGGATTG GCCAAAGGAC CCAAAGGTAT GTTTCGAATG ATACTAACAT
    ACTTATAAGT TAAAAGTCT GGGCCCTAAC CGTTTTCCTG GGTTCCTATA CAAAGCTTAC TATGATTGTA

synth. intron A                               MCS
=====
141 AACATAGAAC AATTTTCAGGA GGACCCCTGG CTAGCGTCCT GCTGGGATTA CACATGGCAT GGATGAACATA
    TTGTATCTTG TAAAAGTCTT CCTGGGAACC GATCGCAGGA CGACCCCTAAT GTGTACCGTA CCTACTTGAT

=====
unc-54 3' UTR
=====
211 TACAAATAGG GCGGGCCGAG CTCGCGATCG GCCGCTGTCA TCAGATCGCC ATCTCGCGCC CGTGCCCTCTG
    ATGTTTATCC CGGCCGGCTC GAGGCGTAGC CGGCGACAGT AGTCTAGCGG TAGAGCGCGG GCACGGAGAC

=====
unc-54 3' UTR
=====
281 ACTTCTAAGT CCAATTACTC TTCAACATCC CTACATGCTC TTTCTCCCTG TGCTCCACCC CCCTATTTTT
    TGAAGATTCA GGTAAATGAG AAGTTGTAGG GATGTACGAG AAAGAGGGAC ACGAGGGTGG GGGATAAAAA

=====
unc-54 3' UTR
=====
351 GTTATTATCA AAAAACTTC TTCTTAATTT CTTGTGTTTT TAGCTTCTTT TAAGTACCT CTAACAATGA
    CAATAATAGT TTTTTGAAG AAGAATTAAG GAAACAAAAA ATCGAAGAAA ATTCAAGTGA GATTGTACT

=====
unc-54 3' UTR
=====
421 AATTGTGTAG ATTCAAAAAT AGAATTAATT CGTAATAAAA AGTCGAAAAA AATTGTGCTC CCTCCCCCA
    TTAACACATC TAAGTTTTTA TCTTAATTAA GCATTATTTT TCAGCTTTTT TTAACACGAG GATGGGGGT

=====
unc-54 3' UTR
=====
491 TTAATAATTA TTCTATCCCA AAATCTACAC AATGTTCTGT GTACACTTCT TATGTTTTTT TTAATTCTGA
    AATTATTATT AAGATAGGGT TTAGATGTG TTACAAGACA CATGTGAAGA ATACAAAAAA AATGAAGACT

=====
unc-54 3' UTR
=====
561 TAAATTTTTT TTGAACATC ATAGAAAAAA CCGCACACAA AATACCTTAT CATATGTTAC GTTTCAGTTT
    ATTTAAAAAA AACTTTGTAG TATCTTTTTT GCGGTGTGTT TTATGGAATA GTATACAATG CAAAGTCAAA

unc-54 3' UTR
=====
631 ATGACCGCAA TTTTATTTC TTCGCACGTC TGGGCTCTC ATGACGTCAA ATCATGCTCA TCGTGAAAAA
    TACTGGCGTT AAAAATAAAG AAGCGTGAG ACCCGGAGAG TACTGCAGTT TAGTACGAGT AGCACTTTTT

unc-54 3' UTR
=====
701 GTTTTGGAGT ATTTTGGAA TTTTCAATC PAGTGAAAGT TTATGAAAT AATTTTCCTG CTTTGTCTTT
    CAAAACCTCA TAAAAACCTT AAAAAGTTAG TTCACTTTCA AATACTTTAA TAAAAGGAC GAAAACGAAA

unc-54 3' UTR
=====
771 TTGGGGGTTT CCCCTATTGT TTGTCAAGAG TTTCGAGGAC GGCCTTTTTT TTGCTAAAAA CACAAGTATT
    AACCCCAAAA GGGGATAACA AACAGTCTC AAAGCTCCTG CCGCAAAAAG AACGATTTTA GTGTTCATAA

unc-54 3' UTR
=====
841 GATGAGCAGC ATGCAAGAAA GATCGGAAGA AGGTTTGGGT TTGAGGCTCA GTGGAAGGTG AGTAGAAGTT
    CTACTCGTGC TACGTTCTTT CTAGCTTCTT TCCAAACCCA AACTCGGAGT CACCTCCAC TCATCTTCAA

unc-54 3' UTR
=====
911 GATAATTTGA AAGTGGAGTA GTGTCTATGG GGTTTTGGCC TTAAATGACA GAATACATTC CCAATATACC
    CTATTAACT TTCACCTCAT CACAGATACC CCAAAACGG AATTACTGT CTTATGTAAG GTTTATATGG

unc-54 3' UTR
=====
981 AAACATAACT GTTTCCTACT AGTCGGCCGT ACGGGCCCTT TCGTCTCGCG CGTTTCGGTG ATGACGGTGA
    TTTGTATTGA CAAAGGATGA TCAGCCGGCA TGCCCGGGA AGCAGAGCGC GCAAAGCCAC TACTGCCACT

```

FIG. 5 (CONTINUED 1.)

1051 AAACCTCTGA CACATGCAGC TCCCGGAGAC GGTACACAGT TGTCTGTAAG CGGATGCCGG GAGCAGACAA
TTTGGAGACT GTGTACGTGG AGGGCCTCTG CCAGTGTCTGA ACAGACATTC GCCTACGGCC CTCGCTCTGT

1121 GCCCGTCAGG GCGCGTCAGC GGGTGTGGC GGGTGTCCGG GCTGGCTTAA CTATGCGGCA TCAGAGCAGA
CGGGCAGTCC GCGCGAGTCG CCCACAACCG CCCACAGCCC CGACCGAATT GATACGCCGT AGTCTCTGTCT

1191 TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA AGGAGAAAAT ACCGCAATCAG
AACATGACTC TCACGTGGTA TACGCCACAC TTTATGGCGT GTCTACGCAT TCCTCTTTTA TGGCGTAGTC

1261 GCGGCCCTTAA GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC
CGCCGGAATT CCCGGAGCAC TATGCGGATA AAAATATCCA ATTACAGTAC TATTATTACC AAAGAATCTG

1331 GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT
CAGTCCACCG TGAAAAGCCC CTTTACACGC GCCTTGGGGA TAAACAAATA AAAAGATTTA TGTAGTTTA

amp
=====

1401 ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA
TACATAGGCG AGTACTCTGT TATTGGGACT ATTTACGAAG TTATTATAAC TTTTTCCTTC TCATACTCAT

amp
=====

1471 TTCAACATTT CCGTGTCCGC CTTATTCCCT TTTTTCGGC ATTTTGCCTT CCTGTTTTG CTCACCCAGA
AAGTTGTAAG GGCACAGCGG GAATAAGGGA AAAAACGCGG TAAACAGGAA GGACAAAAAC GAGTGGGTCT

amp
=====

1541 AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC
TTGCGACCAC TTTCAATTTT TACGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT TGACCTAGAG

amp
=====

1611 AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTAAAGTTC
TTGTCGCCAT TCTAGGAAGT CTCAAAGCG GGGCTTCTTG CAAAGGTTA CTACTCGTGA AAATTTCAAG

amp
=====

1681 TGCTATGTGG CGCGGTATTA TCCCGTATTG ACGCGGGGCA AGAGCAACTC GGTGCGCGCA TACACTATTTC
ACGATACACC GCGCCATAAT AGGGCATAAC TGCGGCCCGT TCTCGTTGAG CCAGCGGCGT ATGTGATAAG

amp
=====

1751 TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA
AGTCTTACTG AACCAACTCA TGAGTGGTCA GTGTCTTTT GTAGAATGCC TACCGTACTG TCATTCTCTT

amp
=====

1821 TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC
AATACGTCAC GACGGTATTG GTACTACTA TTGTGACGCC GGTGAATGA AGACTGTTGC TAGCCTCTCTG

amp
=====

1891 CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACCTCG CTTGATCGTT GGGAAACCGGA
GCTTCTCTGA TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG GAAC TAGCAA CCCTTGGCCT

amp
=====

1961 GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC
CGACTTACTT CGGTATGTTT TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCGTTG TTGCAACGCG

FIG. 5 (CONTINUED 2.)

amp
=====

2031 AAACTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT AGACTGGATG GAGGCGGATA
TTTGATAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTAAATTA TCTGACCTAC CTCGCGCTAT

amp
=====

2101 AAGTTGCAGG ACCACTTCTG CGCTCGGCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG
TTCAACGTCC TGGTGAAGAC GCGAGCCGGG AAGGCCGACC GACCAAAATA CGACTATTTA GACCTCGGCC

amp
=====

2171 TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC
ACTCGCACCC AGAGCGCCAT AGTAACGTCTG TGACCCCGGT CTACCATTCG GGAGGGCATA GCATCAATAG

amp
=====

2241 TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA
ATGTGCTGCC CCTCAGTCCG TTGATACCTA CTTGCTTTAT CTGTCTAGCG ACTCTATCCA CGGAGTGACT

amp
=====

2311 TTAAGCATTG GTAACGTGCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCATTTTAA
AATTCGTAAC CATTGACAGT CTGGTTCAAA TGAGTATATA TGAATCTAA CTAAATTTTG AAGTAAAAAT

2381 ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTCG
TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT AGGGAATTGC ACTCAAAAGC

2451 TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA
AAGGTGACTC GCAGTCTGGG GCATCTTTTC TAGTTTCCTA GAAGAAGTCT AGGAAAAAAA GACGCGCATT

2521 TCTGTGCTT GCAAAACAAA AAACCACCGC TACCAGCGGT GGTGTGTTG CCGGATCAAG AGCTACCAAC
AGACGACGAA CGTTTGTGTT TTTGGTGGCG ATGGTCGCCA CCAACAAAC GGCCTAGTTC TCGATGGTTG

2591 TCTTTTCCG AAGGTAACG GCTTCAGCAG AGCGCAGATA CCAATACTG TCCTTCTAGT GTAGCCGTAG
AGAAAAAGG TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTTTATGAC AGGAAGATCA CATCGGCATC

2661 TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCACTGG
AATCCGGTGG TGAAGTTCCT GAGACATCGT GCGGATGTA TGGAGCGAGA CGATTAGGAC AATGGTCACC

2731 CTGCTGCCAG TGGCGATAAG TCGTGCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA
GACGACGGTC ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCGTCT ATCAATGGCC TATTCCGCGT

2801 GCGGTCGGGC TGAACGGGG GTTCGTGCAC ACAGCCGAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA
CGCCAGCCCG ACTTGCCCC CAAGCACGTG TGTGCGGTG AACCTCGCTT GCTGGATGTG GCTTGACTCT

2871 TACCTACAGC GTGAGCATTG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GCGGACAGG TATCCGGTAA
ATGGATGTCG CACTCGTAAC TCTTCGCGG TCGGAAGGGC TTCCCTCTTT CCGCCTGTCC ATAGGCCATT

2941 GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGAAAC GCCTGGTATC TTTATAGTCC
CGCCGTCCCA GCCTTGTCCT CTCGCGTGTG CCTCGAAGG TCCCCCTTTG CGGACCATAG AAATATCAGG

3011 TGTGCGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTGT TGATGCTCGT CAGGGGGGCG GAGCCTATGG
ACAGCCCAAA GCGGTGGAGA CTGAACCTCG AGCTAAAAAC ACTACGAGCA GTCCCCCGCG CTCGGATACC

3081 AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC ATGTTCTTTC
TTTTTGCGGT CGTTGCGCCG GAAAAATGCC AAGGACCGGA AAACGACCGG AAAACGAGTG TACAAGAAAG

3151 CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGAG
GACGCAATAG GGGACTAAGA CACCTATTGG CATAATGGCG GAAACTCACT CGACTATGGC GAGCGGCGTC

3221 CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC
GGCTTGCTGG CTCGCGTCGC TCAGTCACTC GTCCTTCGC CTTCTCGCGG GTTATGCGTT TGGCGGAGAG

3291 CCGCGCGGTT GCGCGATTCA TTAATGCAGC TGGCACGACA GGTTCCTCGA CTGGAAGCG GCGAGTGAGC
GGGCGGCGAA CCGGCTAAGT AATTACGTCG ACCGTGCTGT CCAAAGGGCT GACCTTTCGC CCGTCACTCG

3361 GCAACGCAAT TAATGTGAGT TAGCTCACTC ATTAGGCACC CAGGCTTTA CACTTTATGC TTCGGGCTCG
CGTTGCGTTA ATTACACTCA ATCGAGTGAG TAATCCGTGG GTTCCGAAAT GTGAAATACG AAGGCCGAGC

3431 TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTACACA GGAACAGCT ATGACCATGA TTACGCCA
ATACAACACA CCTTAACACT CGCCTATTGT TAAAGTGTGT CTTTGTCTGA TACTGGTACT AATGCGGT

FIG. 6.

```

===== T7 promoter ===== Outron
1  AGCTTGGCGC CTAATACGAC TCACATATAGG GCTGCAGGTC GACTCTAGAT TACAACAAAT TATACTTATT
   TCGARCCGCG GATTATGCTG AGTGATATCC CGACGTCCAG CTGAGATCTA ATGTTGATTA ATATGAATAA

Outron
71  TGAATATTCA AATTTTCAGA CCCGGGATTG GCCAAAGGAC CCAAAGGTAT GTTTCGAATG ATACTAACAT
   ACTTATAAGT TTAAGATCTT GGGCCCTAAC CGGTTTCCTG GGTTCCTATA CAAAGCTTAC TATGATTGTA

synth. intron A
141 AACATAGAAC ATTTTCAGGA GGACCCTTGG CTAGCGTCGA CGGTACCATG GGGCGCGCCA TGAGTAAAGG
   TTGTATCTTG TAAAAGTCTT CCTGGGAACC GATCGCAGCT GCCATGGTAC CCCGCGCGGT ACTCATTTC

GFP with introns
211 AGAAGAACTT TTCCTGGAG TTGTCCCAAT TCTTGTGAA TTAGATGGTG ATGTTAATGG GCACAAATTT
   TCTTCTGAA AAGTGACCTC AACAGGGTTA AGAACAACCT AATCTACCAC TACAATTACC CGTGTTTAAA

GFP with introns
281 TCTGTCACTG GAGAGGGTGA AGGTGATGCA ACATACGGAA AACTTACCCT TAAATTTATT TGCCTACTG
   AGACAGTCAC CTCTCCCACT TCCACTACGT TGTATGCCTT TTGAATGGGA ATTTAAATAA ACGTGATGAC

GFP with introns
351 GAAACTACC TGTTCCATGG GTAAGTTTAA ACATATATAT ACTAACTAAC CCTGATTATT TAAATTTTCA
   CTTTGATGG ACAAGGTACC CATTCAAATT TGTATATATA TGATTGATTG GGACTAATAA ATTTAAAGT

GFP with introns
421 GCCAACACTT GTCCTACTT TCTGTTATGG TGTCAATGC TTCTCGAGAT ACCCAGATCA TATGAAACGG
   CGGTGTGAA CAGTGATGAA AGACAATACC ACAAGTTACG AAGAGCTCTA TGGGTCTAGT ATACTTTGCC

GFP with introns
491 CATGACTTTT TCAAGAGTGC CATGCCCGAA GGTATGTAC AGGAAAGAAC TATATTTTTC AAAGATGACG
   GTACTGAAA AGTTCTCAGC GTACGGGCTT CCAATACATG TCCTTTCTTG ATATAAAAAG TTTCTACTGC

GFP with introns
561 GGAACACAA GACACGTAAG TTTAAACAGT TCGGTACTAA CTAACCATA CATTATTAAT TTTCAGGTGC
   CCTTGATGTT CTGTGCATTC AAATTTGTCA AGCCATGATT GATTGGTATG TATAAATTA AAAGTCCACG

GFP with introns
631 TGAAGTCAAG TTTGAAGGTG ATACCCTTGT TAATAGAATC GAGTTAAAAG GTATTGATTT TAAAGAAGAT
   ACTTCAGTTC AAACCTCCAC TATGGGAACA ATTATCTTAG CTCAATTTTC CATAACTAAA ATTCTTCTA

GFP with introns
701 GGAAACATTC TTGGACACAA ATTGGAATAC AACTATACT CACACATGT ATACATCATG GCAGACAAAC
   CCTTGTAAG AACCTGTGTT TAACCTTATG TTGATATTGA GTGTGTTACA TATGTAGTAC CGTCTGTTG

GFP with introns
771 AAAAGAATGG AATCAAAGTT GTAAGTTTAA ACTTGGACTT ACTAACTAAC GGATTATATT TAAATTTTCA
   TTTTCTTACC TTAGTTTCAA CATTCAAATT TGAACCTGAA TGATTGATTG CCTAATATAA ATTTAAAGT

GFP with introns
841 GAACITCAAA ATTAGACACA ACATTGAAGA TGAAGCGTT CAACTAGCAG ACCATTATCA ACAAATACT
   CTTGAAGTTT TAATCTGTGT TGTAACTTCT ACCTTCGCAA GTTGATCGTC TGGTAATAGT TGTTTTATGA

GFP with introns
911 CCAATTGCGG ATGGCCCTGT CTTTTCACCA GACAACCAT ACCTGTCCAC ACAATCTGCC CTTTCGAAAG
   GGTTAACCGC TACCGGGACA GGAATAGTGT CTGTTGGTAA TGGACAGGTG TGTTAGACGG GAAAGCTTTC

GFP with introns
981 ATCCCAACGA AAAGAGAGAC CACATGGTCC TTCTTGAGTT TGTAAACAGCT GCTGGGATTA CACATGGCAT
   TAGGGTTGCT TTTCTCTCTG GTGTACCAGG AAGAACTCAA ACATTTGCGA CGACCCTAAT GTGTACCGTA

```

FIG. 6 (CONTINUED 1.)

GFP with introns

unc-54 3' UTR

1051 GGATGAACTA TACAAATAGG GCCGGGCCGAG CTCGCGATCG GCCGCTGTCA TCAGATCGCC ATCTCGCGCC
CCTACTTGAT ATGTTTATCC CGGCCGGCTC GAGCGGTAGC CGCGGACAGT AGCTAGCGG TAGAGCGCGG

unc-54 3' UTR

1121 CGTGCCTCTG ACTTCTAAGT CCAATTACTC TTCAACATCC CTACATGCTC TTTCTCCCTG TGCTCCCACC
GCACGGAGAC TGAAGATTCA GGTAAATGAG AAGTTGTAGG GATGTACGAG AAAGAGGGAC ACGAGGGTGG

unc-54 3' UTR

1191 CCCTATTTTT GTTATTATCA AAAAACTTC TTCTTAATTT CTTTGTTTTT TAGCTTCTTT TAAGTCACCT
GGGATAAAAA CAATAATAGT TTTTGTGAAG AAGAATTAAA GAAACAAAAA ATCGAAGAAA ATTCAGTGGG

unc-54 3' UTR

1261 CTAACAAATGA AATTGTGTAG ATTCAAAAT AGAATTAAT CGTAATAAAA AGTCGAAAAA AATTGTGCTC
GATTGTACT TTAACACATC TAAGTTTTTA TCTTAATTAA GCATTATTTT TCAGCTTTTT TTAACACGAG

unc-54 3' UTR

1331 CCTCCCCCA TTAATAATAA TTCTATCCCA AAATCTACAC AATGTTCTGT GTACACTTCT TATGTTTTT
GGAGGGGGT AATTATTATT AAGATAGGGT TTAGATGTG TTACAAGACA CATGTGAAGA ATACAAAAA

unc-54 3' UTR

1401 TTACTTCTGA TAAATTTTT TTGAAACATC ATAGAAAAA CCGCACACAA AATACCTTAT CATATGTTAC
AATGAAGACT ATTTAAAAA AACTTTGTAG TATCTTTTTT GCGGTGTGTT TTATGGAATA GTATACAATG

unc-54 3' UTR

1471 GTTTCAGTTT ATGACCGCAA TTTTATTTT TTCCGACGTC TGGGCTCTC ATGACGTCAA ATCATGCTCA
CAAAGTCAAA TACTGGCGTT AAAAATAAAG AAGCGTGCAG ACCCGGAGAG TACTGCAGT TAGTACGAGT

unc-54 3' UTR

1541 TCGTGA AAAA GTTTTGGAGT ATTTTGGAA TTTTCAATC AAGTGAAAGT TTATGAAAT AATTTTCTGT
AGCATT TTTT CAAAACCTCA TAAAAACCTT AAAAAGTTAG TTCATTTC AATACITTA TAAAAGGAC

unc-54 3' UTR

1611 CTTTTGCTTT TGGGGGTTT CCCCTATTGT TTGTCAAGAG TTTGAGGAC GCGGTTTTT TTGCTAAAAAT
GAAAACGAAA AACCCCAAAA GGGGATAACA AACAGTTCTC AAAGCTCCTG CCGCAAAAAG AACGATTTTA

unc-54 3' UTR

1681 CACAAGTATT GATGAGCACG ATGCAAGAAA GATCGGAAGA AGGTTTGGGT TTGAGGCTCA GTGGAAGGTG
GTGTTCAATA TACTCGTGC TACGTCTTT CTAGCCTTCT TCCAAACCCA AACTCCGAGT CACCTTCCAC

unc-54 3' UTR

1751 AGTAGAAGTT GATAATTTGA AAGTGGAGTA GTGTCTATGG GGTTTTGGC TTAAATGACA GAATACATTC
TCATCTTCAA CTATTAAACT TTCACCTCAT CACAGATACC CAAAAACGG AATTACTGT CTTATGTAAG

unc-54 3' UTR

1821 CCAATATACC AAACATAACT GTTTCCTACT AGTCGGCCGT ACGGGCCCTT TCGTCTCGCG GTTTCGGTG
GGTTATATGG TTTGTATTGA CAAAGGATGA TCAGCCGGCA TGCCCGGGAA AGCAGAGCGC GCAAGCCAC

1891 ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC GGTACAGCT TGTCTGTAAG CGGATGCCGG
TACTGCCACT TTTGGAGACT GTGTACGTCG AGGGCTCTG CCAGTGTCTG ACAGACATTC GCCTACGGCC

1961 GAGCAGACAA GCCCGTCAGG GCGCGTCAGC GGTGTGTTGG GGTGTCTGGG GCTGGCTTAA CTATCGGGCA
CTCGTCTGTT CGGGCAGTCC CGCGCAGTCC CCCACAACCG CCCACAGCCC CGACCGAATT GATACGCCGT

2031 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA AGGAGAAAAT
AGTCTCGTCT AACATGACTC TCACGTGGTA TACGCCACAC TTTATGGCGT GTCTACGCAT TCCTTTTTTA

2101 ACCGCATCAG GCGGCCTTAA GGGCTCTGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG
TGCGGTAGTC CGCCGGAATT CCCGGAGCAC TATGCGGATA AAAATATCCA ATTACAGTAC TATTATTACC

FIG. 6 (CONTINUED 2.)

2171 TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTCTAAAT
AAAGAATCTG CAGTCCACCG TGAAAGGCC CTTTACACGC GCCTTGGGGA TAAACAATA AAAAGATTTA

2241 ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG AAAAGGAAG
TGTAAGTTA TACATAGGCG AGTACTCTGT TATTGGGACT ATTTACGAAG TTATTATAAC TTTTCCTTC

=====
amp
=====

2311 AGTATGAGTA TTCAACATTT CCGTGTCCGC CTTATTCCCT TTTTTCGGC ATTTTGCTT CCGTTTTTG
TCATACTCAT AAGTTGTAAA GGCACAGCGG GAATAAGGGA AAAAACGCCG TAAAACGGAA GGACAAAAAC

=====
amp
=====

2381 CTCACCCAGA AACGCTGGTG AAAGTAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA
GAGTGGGTCT TTGCGACCAC TTTATTTC TACGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT

=====
amp
=====

2451 ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT
TGACCTAGAG TTGTCGCCAT TCTAGGAACT CTCAAAGCG GGGCTTCTTG CAAAAGGTTA CTACTCGTGA

=====
amp
=====

2521 TTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCCCGGCA
AAATTTCAAG ACGATACACC GCGCCATAAT AGGCATAAC TCGGGCCCGT TCTCGTTGAG CCAGCGGCGT

=====
amp
=====

2591 TACACTATTC TCAGAAAGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC
ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGGTCA GTGTCTTTC GTAGAATGCC TACCGTACTG

=====
amp
=====

2661 AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG
TCATTCTCTT AATACGTCAC GACGGTATTG GTACTCACTA TTGTGACGCC GGTGAATGA AGACTGTTGC

=====
amp
=====

2731 ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC CTGATCGTT
TAGCCTCCTG GCTTCTCGA TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG GAACTAGCAA

=====
amp
=====

2801 GGGAAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGGCAAC
CCCTTGGCCT CGACTTACTT CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCGTTG

=====
amp
=====

2871 AACGTTGCGC AAATATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT AGACTGGATG
TTGCAACGCG TTTGATAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA TCTGACCTAC

=====
amp
=====

2941 GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT
CTCCGCCTAT TTCAACGTCC TGGTGAAGAC GCGAGCCGGG AAGGCCGACC GACCAATAA CGACTATTTA

=====
amp
=====

3011 CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT
GACCTCGGCC ACTCGACCC AGAGCGCCAT AGTAACGTCG TGACCCCGGT CTACCATTCG GGAGGGCATA

=====
amp
=====

3081 CGTAGTTATC TACACGACGG GGAGTCAGG AACTATGGAT GAACGAATA GACAGATCGC TGAGATAGGT
GCATCAATAG ATGTGCTGCC CCTCAGTCCG TTGATACCTA CTGCTTTTAT CTGTCTAGCG ACTCTATCCA

FIG. 6(CONTINUED 3.)

amp
=====

3151 GCCTCACTGA TTAAGCATTG GTAACGTGCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAAC
CGGAGTGACT AATTCGTAAC CATTGACAGT CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAATTTTGG

3221 TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA TCCCTTAACG
AAGTAAAAAT TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT AGGGAATTGC

3291 TGAGTTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT
ACTCAAAAGC AAGGTGACTC GCAGTCTGGG GCATCTTTTC TAGTTTCCTA GAAGAACTCT AGGAAAAAAA

3361 CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTTTG CCGGATCAAG
GACGCGCATT AGACGACGAA CGTTTGTTTT TTTGGTGCGG ATGGTCGCCA CCAAACAAAC GGCCTAGTTC

3431 AGCTACCAAC TCTTTTCCG AAGGTAACG GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT
TCGATGGTTG AGAAAAAGGC TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTTTATGAC AGGAAGATCA

3501 GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG
CATCGGCATC AATCCGGTGG TGAAGTTCTT GAGACATCGT GGCGGATGTA TGGAGCGAGA CGATTAGGAC

3571 TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG
AATGGTCACC GACGACGGTC ACCGCTATTG AGCAGAGAAT GGCCCAACCT GAGTTCGTCT ATCAATGGCC

3641 ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCAGC TTGGAGCGAA CGACCTACAC
TATTCGCGCT CGCCAGCCCG ACTTGCCCC CAAGCACGTG TGTGGGTCG AACCTCGCTT GCTGGATGTG

3711 CGAACTGAGA TACCTACAGC GTGAGCATTG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG
GCTTGACTCT ATGGATGTCG CACTCGTAAC TCTTTCGCGG TCGGAAGGGC TTCCCTCTTT CGCCTGTCC

3781 TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAC GCCTGGTATC
ATAGGCCATT CGCCGTCCA GCCTGTCTCT CTCGCGTCT CCCTCGAAGG TCCCCCTTTG CGGACCATAG

3851 TTTATAGTCC TGTGCGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTGG TGATGCTCGT CAGGGGGGCG
AAATATCAGG ACAGCCCAA GCGGTGGAGA CTGAATCGC AGCTAAAAAC ACTACGAGCA GTCCCCCGC

3921 GAGCCTATGG AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC
CTCGGATACC TTTTGCCTG CGTTGCGCGG GAAAAATGCC AAGGACCGGA AAACGACCGG AAAACGAGTG

3991 ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG
TACAAGAAAG GACGCAATAG GGGACTAAGA CACCTATTGG CATAATGGCG GAAACTCACT CGACTATGGC

4061 CTCGCCGCGC CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA
GAGCGGCGTC GGCTTGCTGG CTCGCGTCGC TCAGTCACTC GCTCCTTCGC CTTCTGCGG GTTATGCGTT

4131 ACCGCCCTCT CCCGCGCGTT GGCCGATTCA TTAATGCAGC TGGCAGACA GGTTTCCCGA CTGGAAAGCG
TGGCGGAGAG GGGCGCGCAA CCGGCTAAGT AATTACGTCG ACCGTGCTGT CCAAAGGGCT GACCTTTCGC

4201 GGCACTGAGC GCAACGCAAT TAATGTGAGT TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC
CCGTCACTCG CGTTGCGTTA ATTACACTCA ATCGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG

4271 TTCCGGCTCG TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTACACA GGAAACAGCT ATGACCATGA
AAGGCCGAGC ATACAACACA CCTTAACACT CGCCTATTGT TAAAGTGTGT CTTTGTGCGA TACTGGTACT

4341 TTACGCCA
AATGCGGT

SEQUENCE LISTING

<110> DEVGEN NV

<120> GENE EXPRESSION SYSTEM

<130> SCB/55177/001

<140>

<141>

<160> 5

<170> PatentIn Ver. 2.0

<210> 1

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligonucleotide o-GN59

<400> 1

ctagattaca actaattata cttatttgaa tattcaaatt ttcagac

47

<210> 2

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligonucleotide o-GN60

<400> 2

ccgggtctga aaatttgaat attcaaataa gtataattag ttgtaat

47

<210> 3

<211> 3498

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid
pDW3123

<400> 3

```

agcttggcgc ctaatacgac tcactatagg gctgcaggtc gactctagat tacaactaat 60
tatacttatt tgaatattca aattttcaga ccgggattg gccaaaggac ccaaaggat 120
gtttcgaatg atactaacat aacatagaac attttcagga ggacccttgg ctagcgtcct 180
gctgggatta cacatggcat ggatgaacta tacaaatagg gccggccgag ctccgcatcg 240
gccgctgtca tcagatcgcc atctcgcgcc cgtgcctctg acttctaagt ccaattactc 300
ttcaacatcc ctacatgctc tttctccctg tgctccacc ccctattttt gttattatca 360
aaaaaacttc ttcttaattt ctttgtttt tagcttcttt taagtcacct ctaacaatga 420
aattgtgtag attcaaaaat agaattaatt cgtaataaaa agtcgaaaaa aattgtgctc 480
cctcccccca ttaataataa ttctatccca aaatctacac aatgttctgt gtacacttct 540
tatgtttttt ttacttctga taaatttttt ttgaaacatc atagaaaaaa ccgcacacaa 600
aataccttat catatgttac gtttcagttt atgaccgcaa tttttatttc ttcgcacgtc 660
tgggcctctc atgacgtcaa atcatgtca tcgtgaaaaa gttttggagt atttttggaa 720

```

```

tttttcaatc aagtgaaggt ttatgaaatt aattttcctg cttttgcttt ttggggggttt 780
cccctattgt ttgtcaagag ttctgaggac ggcgttttct ttgctaaaaat cacaagtatt 840
gatgagcacg atgcaagaaa gatcggaaga aggttttgggt ttgaggctca gtggaagggtg 900
agtagaaggt gataatttga aagtggagta gtgtctatgg ggtttttgcc ttaaatgaca 960
gaatacattc ccaatatacc aaacataact gtttcctact agtcggccgt acggggccctt 1020
tcgtctcgcg cgtttcgggt atgacggtga aaacctctga cacatgcagc tcccgagagc 1080
ggtcacaagt tgcctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc 1140
gggtgttggc ggggtcggg gctggcttaa ctatgcggca tcagagcaga ttgtactgag 1200
agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag 1260
gcggccctaa gggcctcgtg atacgcctat ttttataggt taatgtcatg ataataatgg 1320
tttcttagag tgcagtgggt acttttcggg gaaatgtgcg cggaaccctt atttgtttat 1380
ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc 1440
aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc ctattccctt 1500
tttttgccgc attttgcctt cctgtttttg ctcaccacaga aacgctggtg aaagtaaaag 1560
atgctgaaga tgcagtgggt gcacgagtgg gttacatcga actggatctc aacagcggta 1620
agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc 1680
tgctatgttg cgcggtatta tcccgatttg acgcccggca agagcaactc ggtcgcgcga 1740
tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg 1800
atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg 1860
ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca 1920
tgggggatac tgtaactcgc ctgtatcgtt gggaaccgga gctgaatgaa gccataccaa 1980
acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgcg aaactattaa 2040
ctggcgaaact acttactcta gcttcccggc aacaattaat agactggatg gaggcggata 2100
aagttgcagg accacttctg cgtcggccct ttccggctgg ctggtttatt gctgataaat 2160
ctggagccgg tgagcgtggg tctcgcggtg tcatgtcagc actggggcca gatggttagc 2220
cctcccgtag cgtagtattc tacacgacgg ggagtcaggg aactatggat gaacgaaata 2280
gacagatcgc tgagataggt gctcactga ttaagcattg gtaactgtca gaccaagttt 2340
actcatatat actttagatt gatttaaaac ttcattttta atttaaaagg atctaggtga 2400
agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg ttccactgag 2460
cgtcagaccc cgtagaaaaa atcaaaggat cttcttgaga tccctttttt ctgcgcgtaa 2520
tctgctgctt gcaaaacaaa aaaccacgc taccagcggg ggtttgtttg ccggatcaag 2580
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg 2640
tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat 2700
acctcgctct gctaatectg ttaccagtgg ctgctgccag tggcgataag tctgttctta 2760
ccgggttgga ctcaagacga tagttaccgg ataagcgca gcggtcgggc tgaacggggg 2820
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc 2880
gtgagcattg agaaagcgcc acgcttcccg aaggagagaa ggcggacagg tatccggtaa 2940
gcggcagggg cggaacagga gagcgcacga gggagcttcc agggggaac gcttggtatc 3000
tttatagtc tgcctgggtt cgccacctct gacttgagcg tcgatttttg tgatgctcgt 3060
cagggggggg gagcctatgg aaaaacgcca gcaacggcgc ctttttacgg ttccctggcct 3120
tttgctggcc ttttgctcac atgttcttcc ctgcgttatc cctgattct gtggataacc 3180
gtattaccgc ctttgagtga gctgataccg ctgcgcgag ccgaacgacc gagcgagcgc 3240
agtcagtgag cgaggaagcg gaagagcgcc caatacgaac accgcctctc ccgcgcggtt 3300
ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaagcg ggcagtgagc 3360
gcaacgcaat taatgtgagt tagctcactc attagccacc ccaggcttta cactttatgc 3420
ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaacagct 3480
atgaccatga ttacgcca
3498

```

<210> 4

<211> 4348

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid
pDW3124

<400> 4

```

agcttgccgc ctaatacgac tcaactatagg gctgcaggtc gactctagat tacaactaat 60
tatacttatt tgaatattca aattttcaga cccgggattg gccaaaggac ccaaaggat 120
gtttcgaatg atactaacat aacatagaac attttcagga ggacccttgg ctagcgtcga 180

```

cggtaccatg gggcgcgcca tgagtaaagg agaagaactt ttacttgag ttgtcccaat 240
 tcttggtgaa ttatagtggt atgttaatgg gcacaaattt tctgtcagtg gagagggtga 300
 aggtgatgca acatacggaa aacttaccct taaatttatt tgcactactg gaaaactacc 360
 tgttccatgg gtaagtttaa acatataat actaactaac cctgattatt taaattttca 420
 gccaacactt gtcactactt tctgttatgg tgttcaatgc ttctcgagat acccagatca 480
 tatgaaacgg catgactttt tcaagagtgc catgcccgaa gggtatgtac aggaagaac 540
 tataattttt aaagatgacg ggaactacaa gacacgtaag tttaaacagt tccgtactaa 600
 ctaaccatac atattttaa tttcaggtgc tgaagtcaag ttgaagggtg atacccttgt 660
 taatagaatc gagttaaaag gtattgattt taaagaagat ggaacattc ttggacacaa 720
 attggaatc acataaact cacacaatgt atacatcatg gcagacaaac aaaagaatgg 780
 aatcaagtt gtaagtttaa acttggactt actaactaac ggattatatt taaattttca 840
 gaacttcaaa attagacaca acattgaaga tgggaagcgt caactagcag accattatca 900
 acaaaatata ccaattggcg atggccctgt ccttttacca gacaaccatt acctgtccac 960
 acaatctgcc ctttcgaaag atcccaacga aaagagagac cacatgggtc ttcttgagtt 1020
 tgtaacagct gctgggatta cacatggcat ggatgaacta tacaatagg gccggccgag 1080
 ctccgcatcg gccgctgtca tcagatcgcc atctcgccg cgtgcctctg acttctaagt 1140
 ccaattactc ttcaacatcc ctacatgctc tttctccctg tgcctccacc cctattttt 1200
 gttattatca aaaaaactt tctttaaatt ctttgtttt tagcttcttt taagtacact 1260
 ctaacaatga aattgtgtag attcaaaaat agaattaatt cgtaataaaa agtcgaaaaa 1320
 aattgtgctc cctccccca ttaataataa ttctatccca aaatctacac aatgttctgt 1380
 gtacacttct tatgtttttt ttacttctga taaattttt ttgaaacatc atagaaaaaa 1440
 ccgcacacaa aataccttat catatgttac gtttcagttt atgaccgcaa tttttatttc 1500
 ttcgcacgtc tgggcctctc atgacgtcaa atcatgctca tctgaaaaa gttttggagt 1560
 atttttggaa tttttcaatc aagtgaagtt ttatgaaatt aattttcctg cttttgcttt 1620
 ttggggggtt ccctatttgt ttgtcaagag ttctcaggac ggcgtttttc ttgctaaaat 1680
 cacaagtatt gatgagcagc atgcaagaaa gatcggaaga aggtttgggt ttgaggctca 1740
 gtggaagggt agtagaagtt gataatttga aagtggagta gtgtctatgg gggttttgcc 1800
 ttaaatgaca gaatacatc ccaatatacc aaacataact gtttccctact agtcggccgt 1860
 accggccctt tctctcgcg cgtttcgggt atgacgggtg aaacctctga cacatgcagc 1920
 tcccggagac ggtcacagct tgtctgtaag cggatgccgg gacagacaa gcccgctcag 1980
 gcgcgtcagc ggggtgtggc ggggtgtcgg gctggcctaa ctatgcggca tcagagcaga 2040
 ttgactagag agtgaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 2100
 accgcatcag gccgccttaa gggcctcgtg atacgcctat ttttatagtt taatgtcatg 2160
 ataataatgg tttcttagac gtcagggtggc acttttcggg gaaatgtgcg cggaacccct 2220
 atttgtttat tttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga 2280
 taaatgcttc aataatattg aaaaaggaa agtatgagta ttcaacattt ccgtgtcgcc 2340
 cttattccct tttttgcggc attttgcctt cctgtttttg ctacccaga aacgctgggt 2400
 aaagtaaaag atgctgaaga tcagttgggt gcacgagtg gttacatcga actggatctc 2460
 aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact 2520
 tttaaagttc tgctatgtgg cgcggtatta tcccgattg acgcccggca agagcaactc 2580
 ggtcgcgca tacactatc tcagaatgac ttggttgagt actcaccagt cacagaaaag 2640
 catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat 2700
 aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt 2760
 ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt ggaacccgga gctgaatgaa 2820
 gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgccg 2880
 aaactattaa ctggcgaaact acttactcta gcttcccgcc aacaattaat agactggatg 2940
 gaggcggata aagttgcagg accacttctg cgctcggccc ttccgctggg ctggtttatt 3000
 gctgataaat ctggagccgg tgagcgtggg tctcgggta tcattgcagc actggggcca 3060
 gtggtgaagc cctcccgat cgtagttatc tacacgacgg ggagtcaggc aactatggat 3120
 gaacgaaata gacagatcgc tgagataggt gcctactga ttaagcattg gtaactgtca 3180
 gaccaagttt actcatatat acttttagatt gatttaaaac ttcattttta atttaaaagg 3240
 atctagggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg 3300
 ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccctttttt 3360
 ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgttt 3420
 ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata 3480
 ccaaatactg tcttctagt gtagecgtag ttaggccacc acttcaagaa ctctgtagca 3540
 ccgcctacat acctcgtctt gctaactcgt ttaccagtg ctgctgccag tggcgataag 3600
 tctgtcttta ccgggttggg ctcaagacga tagttaccgg ataaggcgca gcggtcgggc 3660
 tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga 3720
 tacctacagc gtgagcattg agaaagcgcc acgcttcccg aaggagagaa ggcggacagg 3780
 tatccggtaa cgggcagggt cggaaacagga gagcgacga gggagcttcc agggggaaac 3840

```

gctcggtatc tttatagtc tgtcgggttt cgccacctct gacttgagcg tcgatttttg 3900
tgatgctcgt caggggggag gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg 3960
ttcctggcct tttgctggcc ttttgcacac atgttctttc ctgcgttatc ccctgattct 4020
gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgcgcag ccgaacgacc 4080
gagcgagcgg agtcagtgag cgaggaagcg gaagagcgcc caatacgcaa accgcctctc 4140
cccgcgcgtt ggccgattca ttaatgcagc tggcagcaca ggtttccga ctggaaagcg 4200
ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggtttta 4260
cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca 4320
ggaaacagct atgaccatga ttacgcca 4348

```

<210> 5

<211> 9309

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid pGN148

<400> 5

```

atgactgctc caaagaagaa gcgtaaggta ccggtaatga acacgattaa catcgctaag 60
aacgacttct ctgacatcga actggctgct atcccgttca acactctggc tgaccattac 120
ggtgagcggt tagctcggta agtttaaaaca tctagatact aactaacgat taacatttaa 180
attttcagcg aacagttggc ccttgagcat gactcttacg agatgggtga agcagcgttc 240
cgcaagatgt ttgagcgta acttaaaagt ggtgaggttg cgataaacgc tgcccgcaag 300
cctctcatca ctacctact ccctaagatg attgcacgca tcaacgactg gtttgaggaa 360
gtgaaagcta agcgcggcaa gcgcccagca gccttcagtg tcctgcaaga aatcaagccg 420
gaagcgcgtg cgtacatcac cattaagacc actctggctt gcctaaccag tgctgacaat 480
acaaccgttc aggtctgtagc aagcgcaatc ggtcggggcca ttgaggacga ggctcgcttc 540
ggtcgtatcc gtgacctga agctaagcac ttcaagaaaa acgttgagga acaactcaac 600
aagcgcgtag ggcagctcta caagaaagca tttatgcaag ttgtcgaggc tgacatgctc 660
tctaaggggt tactcgtggc cgaggcgtgg tcttcgtggc ataaggaaaga ctctattcat 720
gtaggagtac gctgcatcga gatgctcatt gactcaaccg gactgggttag cttacaccgc 780
caaaatgctg gcgtagtagg tcaagactct gagactatcg aactcgcacc tgaatacgct 840
gaggctatcg caaccctgct aggtgcgctg gctggcatct ctccgatgtt ccaaccttgc 900
gtagttcctc ctaagcgtg gactggcatt actgggtgtg gctattgggc taacggtcgt 960
cgctccttgg cgctggtgct tactcacagt aagaaagcac tgatgcgcta cgaagacgtt 1020
tacatgcctg aggtgtacaa agcgattaac attgcgcaaa acaccgcatg gaaaatcaac 1080
aagaaagtcc tagcggctcg caacgtaatc accaagtggg agcattgtcc ggtcgaggac 1140
atccctgcga ttgagcgtga agaactcccc atgaaaccgg aagacatcga catgaatcct 1200
gaggctctca ccgcgtggaa acgtgctgcc gctgctgtgt accgcaagga caaggctcgc 1260
aagtctcgcc gtatcagcct tgagttcatg cttgagcaag ccaataagtt tgctaaccat 1320
aagcccatct ggttccctta caacatggac tggcgcggtc gtgtttacgc tgtgtcaatg 1380
ttcaaccgca aagctaacga tatgacaaa ggactgctta cgctggcgaa aggtaaacca 1440
atcggttaag aaggttacta ctggctgaaa atccacggtg caaactgtgc ggtgtcgat 1500
aaggttccgt tccctgagcg catcaagttc attgaggaaa accacgagaa catcatggct 1560
tgcgctaagt ctccactgga gaacacttgg tgggctgagc aagattctcc gttctgcttc 1620
cttgcgttct gctttgagta cgctggggta cagcaccacg gcctgagcta taactgctcc 1680
cttccgctgg cgtttgacgg gtcttgcctc ggcatccagc acttctccgc gatgctccga 1740
gatgaggtag gtgtcgcgcg ggttgtaagt ttaactcta tctactaac taacgaagct 1800
tatttaaatt ttcagaactt gcttcctagt gaaaccgttc aggacatcta cgggattgtt 1860
gctaagaaag tcaacgagat tctacaagca gacgcaatca atgggaccga taacgaagta 1920
gttaccgtga ccgatgagaa cactggtgaa atctctgaga aagtcaagct gggcactaag 1980
gcactggctg gtcaatggct ggcttacggt gttactcgca gtgtgactaa gcgttcagtc 2040
atgacgctgg cttacgggtc caaagagttc ggcttccgct aacaagtgtc ggaagatacc 2100
attcagccag ctattgattc cggcaagggt ctgatgttca ctacgcccga tcaggtgct 2160
ggatacatgg ctaagctgat ttgggaatct gtgagcgtga cgggtgtagc tgcggttgaa 2220
gcaatgaact ggcttaagtc tgctgctaag ctgctggctg ctgaggtcaa agataagaag 2280
actggagaga ttcttcgcaa gcgttgcgct gtgcattggg tcaactccga tggtttccct 2340
tgtgtggcagg aatacaagaa gcctattcaa acgcgtttga acctgatgtt cctcggtcag 2400
ttccgcttac agcctaccat taacaccaac aaagatacgc agattgatgc acacaaacag 2460
gagtcgtgta tcgctcctaa ctttgtacac agccaagacg gttagccact tcgtaagact 2520

```


gtagtgtggg cacacgagaa gtacggaatc gaatcttttg cactgattca cgactccttc 2580
 ggtaccattc cggctgacgc tgcgaacctg ttcaaagcag tgcgcgaaac tatgggtgac 2640
 acatatgagt cttgtgatgt actggctgat ttctacgacc agttcgctga ccagttgcac 2700
 gagtctcaat tggacaaaat gccagcactt ccggctaaag gtaacttgaa cctccgtgac 2760
 atcttagagt cggacttcgc gttcgcgtaa gaattccaac tgagcgccgg tcgctaccat 2820
 taccaacttg tctggtgtca aaaataatag gggccgctgt catcagagta agtttaaaact 2880
 gagttctact aactaacgag taataatttaa attttcagca tctcgcgccc gtgcctctga 2940
 cttctaagtc caattactct tcaacatccc tacatgctct ttctccctgt gctcccaccc 3000
 cctatttttg ttattatcaa aaaaacttct tcttaatttc ttgtttttt agcttctttt 3060
 aagtcacctc taacaatgaa attgtgtaga ttcaaaaata gaattaattc gtaataaaaa 3120
 gtcgaaaaaa attgtgctcc ctcccccat taataataat tctatcccaa aatctacaca 3180
 atgttctgtg tacacttctt atgttttttt tacttctgat aaattttttt tgaaacatca 3240
 tagaaaaaac cgcacacaaa ataccttctc atatgttacg ttccagttta tgaccgcaat 3300
 ttttatttct tcgcacgtct gggcctctca tgacgtcaaa tcatgctcat cgtgaaaaag 3360
 ttttgagta tttttggaat ttttcaatca agtgaaagtt tatgaaatta attttcctgc 3420
 ttttgctttt tgggggtttc ccctattggt tgtcaagagt ttccaggacg gcgtttttct 3480
 tgctaaaatc acaagtattg atgagcacga tgcaagaaag atcgggaagaa ggtttgggtt 3540
 tgaggctcag tggaagggtga gtagaagttg ataatttgaa agtggaagtag tgtctatggg 3600
 gtttttgctt taatgacag aatacattcc caatatacca aacataactg tttcctacta 3660
 gtcggccgta cgggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 3720
 acatgcagct ccgagagacg gtcacagctt gtctgtaagc ggatgcggg agcagacaag 3780
 ccgctcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 3840
 cagagcagat tgtactgaga gtgcaccata tgccggttga aataccgcac agatgcgtaa 3900
 ggagaaaata ccgcatcagg cggccttaag ggcctcgtga tacgcctatt tttataggtt 3960
 aatgtcatga taataatggt ttcttagacg tcagggtggca ctttccgggg aaatgtgcgc 4020
 ggaaccctta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 4080
 taaccctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc 4140
 cgtgtcgcctt ttattccctt ttttgccgca ttttgccttc ctgtttttgc tcaccagaa 4200
 acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggt ttacatcgaa 4260
 ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 4320
 atgagcactt ttaagttct gctatgtggc gcggtattat ccggtattga cgcgggca 4380
 gagcaactcg gtcgcgcgat acactattct cagaatgact tgggtgagta ctcaccagtc 4440
 acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 4500
 atgagtata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 4560
 accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcggtt ggaaccggag 4620
 ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgccgtgtagc aatggcaaca 4680
 acgttgccga aactattaac tggcgaacta ctactctag cttcccgga acaattaata 4740
 gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggtggc 4800
 tgggtttattg ctgataaatc tggagccggt gagcgtgggt ctgcggtat cattgcagca 4860
 ctggggccag atggttaagc ctcccgatc gtagttatct acacgacggg gagtcaggca 4920
 actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 4980
 taactgtcag accaagttaa ctcatatata ctttagattg atttaaaact tcatttttaa 5040
 tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt 5100
 gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaaggatc ttcttgagat 5160
 cctttttttc tgcgcgtaat ctgctgcttg caaacaacaa aaccaccgct accagcgtg 5220
 gtttggttgc cggatcaaga gctaccaact ctttttccga aggttaactg cttcagcaga 5280
 gcgcagatc caaatactgt ccttctagt tagccgtagt taggccacca cttcaagaac 5340
 tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgetgccagt 5400
 ggcgataagt cgtgtcttac cgggttgga tcaagacgat agttaccgga taaggcgag 5460
 cggtcgggct gaacggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 5520
 gaactgagat acctacagcg tgagcattga gaaagcgcca cgcttcccga agggagaaag 5580
 gcggacaggt atccggttaag cggcagggtc ggaacaggag agcgacagag ggagcttcca 5640
 gggggaaaac cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 5700
 cgatttttgt gatgctcgtc agggggcggt agcctatgga aaaacgcca caacgcggcc 5760
 tttttacggt tcttggcctt ttgctggcct tttgtcaca tgttctttcc tgcgttatcc 5820
 cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc 5880
 cgaacgaccg agcgacgga gtcagttagc gaggaagcgg aagagcgccc aatacgcaaa 5940
 ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac 6000
 tggaaagcgg cgagttagcg caacgcaatt aatgtgagtt agctcactca ttaggcacc 6060
 caggcttac actttatgct tccggtcgt atgttgtgtg gaattgtgag cggataacaa 6120
 tttcacacag gaaacagcta tgaccatgat tacgccaagc tgtaagttaa aacatgatct 6180

tactaactaa	ctatttctcat	ttaaattttc	agagcttaaa	aatggctgaa	atcactcaca	6240
acgatggata	cgctaacaac	ttggaaatga	aataagcttg	catgcctgca	gagcaaaaaa	6300
atactgcttt	tccttgcaaa	attcgggtgct	ttcttcaaag	agaaactttt	gaagtcggcg	6360
cgagcatttc	cttctttgac	ttctctcttt	ccgcaaaaaa	gcctagcatt	tttattgata	6420
atttgattac	acacactcag	agtctctcga	catgataaag	tgtttcattg	gcactcgccc	6480
taacagtaca	tgacaagggc	ggattattat	cgatcgatat	tgaagacaaa	ctccaaatgt	6540
gtgctcattt	tggagccccg	tgtggggcag	ctgctctcaa	tatattacta	gggagacgag	6600
gagggggacc	ttatcgaaacg	tcgcatgagc	cattctttct	tcctttatgca	ctctcttcac	6660
tctctcacac	attaatcgat	tcatagactc	ccatattcct	tgatgaaggt	gtgggttttt	6720
agcttttttt	cccgatattgt	aaaaggaaga	ggctgacgat	gttaggaaaa	agagaacgga	6780
gccgaaaaaa	catccgtagt	aagtcttctt	tttaagccga	cacttttttag	acagcattcg	6840
ccgctagtgt	tgaagttaa	attttaaaaa	ataaaaaatta	gtttcaattt	tttttaatta	6900
ctaaataggc	aaaagttttt	tcaagaactc	tagaaaaact	agcttaattc	atgggtacta	6960
gaaaaattct	tgttttaaat	ttaatatatta	tcttaagatg	taattacgag	aagctttttt	7020
gaaaattctc	aattaaaaga	atttgccgat	ttagaataaa	agtcttcaga	aatgagtaaa	7080
agctcaaat	agaagtgtgt	ttttaaaaga	aaaacacgaa	aaaagaacac	tatttatctt	7140
ttctctcccg	cgtaaaatta	gttggttga	taatagtgat	ccgctgtcta	tttgactcgc	7200
gctcttcaca	ccgtgcttcc	tctcacttga	ccaacagga	aaaaaaaaaca	tcacgtctga	7260
gacgggtgaat	tgcccttatca	agagcgtcgt	ctctttcacc	cagtaacaaa	aaaaatttgg	7320
tttctttact	ttatatttat	gtaggtcaca	aaaaaaaaagt	gatgcagttt	tggtgggtcgc	7380
ttgtctccac	accacctccg	cctccagcag	cacacaatca	tcttcgtgtg	ttctcgacga	7440
ttccttgtat	gccgcggtcg	tgaatgcacc	acattcgacg	cgcaactaca	caccacactc	7500
actttcgggt	gtattactac	acgtcatcgt	tggtcgtagt	ctcccgtctc	ttcgtcccca	7560
ctcactctc	attattcccc	ttggtgtatt	gatttttttt	aaatggtaca	ccactcctga	7620
cgtttctacc	ttcttggttt	ccgtccattt	agattttatc	tggaaatttt	tttaaaattt	7680
taggcccagag	agttctagt	cttggtctaa	aagtctaggt	cagacataca	tttctctatt	7740
ctcatcaaaa	aaaaagttga	taaagaaaaa	tggttattca	gaaagagtgt	gtctcgttga	7800
aattgtattca	aaaaaaaaat	cccaccctc	gcttggttct	caaaatatga	gatcaacgga	7860
tttttctct	ctcgattcaa	ttttttgctg	cgctctgtct	gccaaagtgt	gtgtgtccga	7920
gcaaaaagatg	agagaattta	caaacagaaa	tgaaaaaaaag	ttggccaaat	aatgaagtgt	7980
tatccgagat	tgatgggaaa	gatattaatg	ttctttacgg	tttgaggagg	agagagagat	8040
agattttcgc	atcaaaactcc	gccttttaca	tgtcttttag	aatctaaaat	agatttttct	8100
catcattttt	aatagaaaaa	cgagaaatta	cagtaatttc	gcaattttct	tgccaaaaat	8160
aca'cgaaatt	tgtgggtctc	gccacgatct	cggtcttagt	ggttcatttg	gtttaaaaagt	8220
ttataaaaatt	tcaaaattcta	gtgtttaatt	tccgcataat	tggacctaaa	atgggttttt	8280
gtcatcattt	tcaacaagaa	atcgtagaaa	tcctgttggt	tcgcaatttt	cttttcaaaa	8340
atacacgaaa	tatatggtaa	tttcccga	tattgagggt	ctcgccacga	tttcagtcac	8400
agtgggcagg	atttatcacg	aaaaaagtcc	gcctagtctc	acatttccgg	aaaaccgaat	8460
ctaaattagt	tttttgatcat	cattttgaac	aaaaaatcga	gacatcccta	tagtttcgca	8520
attttcgtcg	cttttctctc	caaaaatgac	agcttagaat	taaaattcgc	tggaaactgg	8580
accatgatat	cttttctccc	cgtttttcat	tttatttttt	attacactgg	attgactaaa	8640
ggtcaccacc	accgccagtg	tgtgccatat	cacacacaca	cacacacaca	atgtcgagat	8700
tttatgtgtt	atccctgctt	gatttctgtt	cgttgtctct	ctctctctat	tcatcttttg	8760
agccgagaag	ctccagagaa	tggagcacac	aggatcccgc	cgcgcgatgt	cgtcgggaga	8820
tggcgccgcc	tgggaagccg	ccgagagata	tcagggaaga	tcgtctgatt	tctcctcgga	8880
tgccacctca	tctctcgagt	ttctccgcct	gttactccct	gccgaacctg	atatttcccg	8940
ttgtcgtaaa	gagatgtttt	tattttactt	tacaccgggt	cctctctctc	tgccagcaca	9000
gctcagtggt	ggctgtgtgc	tcgggctcct	gccaccggcg	gcctcatctt	cttctctctc	9060
ttctctcctg	ctctcgctta	tcaettcttc	attcattctt	attccttttc	atcatcaaac	9120
tagcattttc	tactttattt	atttttttca	attttcaatt	ttcagataaa	accaaactac	9180
ttgggttaca	gccgtcaaca	gatccccggg	attggccaaa	ggaccctaa	gtatgtttcg	9240
aatgatacta	acataacata	gaacattttc	aggaggaccc	ttgcttggag	ggtaccgagc	9300
tcagaaaaa						9309

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



INTERNATIONAL BUREAU OF PATENT COOPERATION
358, rue de la Confédération
CH-1202 Genève
Tél. 022 733 1111
Fax 022 733 1112
E-mail: info@wipo.int

(43) International Publication Date
22 November 2001 (22.11.2001)

PCT

(10) International Publication Number
WO 01/88114 A3

(51) International Patent Classification: C12N 15/00

(21) International Application Number: PCT/EP01/05794

(22) International Filing Date: 18 May 2001 (18.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0012231.7 19 May 2000 (19.05.2000) GB

(71) Applicant (for all designated States except US): DEV-
GEN NV [BE/BE]; Technologiepark 9, B-9052 Zwij-
naarde (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PLAETINCK,
Geert [BE/BE]; Pontstraat 16, B-9820 Merelbeke (BE).
RENARD, Jean-Pierre [BE/BE]; Peter Benoîtlaan
141, B-9050 Gentbrugge (BE). MORTIER, Katherine
[BE/BE]; Lange Boomgaardstraat 25, B- Gent (BE).
BOGAERT, Thierry [BE/BE]; Wolvendreef 26g, B-8500
Konrijk (BE).

(74) Agent: BOULT WADE TENNANT; Verulam Gardens,
70 Gray's Inn Road, London WC1X 8BT (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW). Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM). European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR). OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted
a patent (Rule 4.17(ii)) for the following designations AE,
AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH,
CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB,
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP,
KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,
ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL,
SZ, TZ, UG, ZW). Eurasian patent (AM, AZ, BY, KG, KZ,
MD, RU, TJ, TM). European patent (AT, BE, CH, CY, DE,
DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR),
OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML,
MR, NE, SN, TD, TG)
- of inventorship (Rule 4.17(iv)) for US only

Published:

- with international search report
— before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments

(88) Date of publication of the international search report:
20 June 2002

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/88114 A3

(54) Title: GENE EXPRESSION SYSTEM

(57) Abstract: There are provided DNA constructs, including replicable cloning vectors and expression vectors, comprising a bac-
tenophage promoter operably linked to an outtron sequence. The expression vectors provided by the invention are useful in the
expression of recombinant polypeptides in host cells or organisms and are particularly useful in expression of recombinant polypep-
tides in nematode worms such as *C. elegans*.

INTERNATIONAL SEARCH REPORT

Int'l Application No

PC1/EP 01/05794

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/63 C12N15/67 C12N15/85 C07K14/005 C12N9/12
A01K67/033

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WIRTZ E ET AL: "Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes" NUCLEIC ACIDS RESEARCH, vol. 22, no. 19, 1994, pages 3887-3894, XP002196041 figure 8	1,9
A	CONRAD R ET AL: "Insertion of part of an intron into the 5' untranslated region of a Caenorhabditis elegans gene converts it into a trans-spliced gene" MOL CELL BIOL, vol. 11, no. 4, April 1991 (1991-04), pages 1921-1926, XP001073410 figure 1B	1-18



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

15 April 2002

Date of mailing of the international search report

29/04/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Lonnoy, O

INTERNATIONAL SEARCH REPORT

Intr. International Application No

PC/EP 01/05794

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 00 01846 A (MORTIER KATHERINE ;DEVGEN NV (BE); BOGAERT THIERRY (BE); PLAETINCK) 13 January 2000 (2000-01-13) cited in the application pGN401	1-18
A	WO 99 37770 A (DEVGEN NV ;BOGAERT THIERRY ANDRE OLIVER E (BE); CRIEKGINGE WIM MARI) 29 July 1999 (1999-07-29) figure 1E	1
A	KRAUSE M: "Transcription and Translation" METHODS CELL BIOL , vol. 48, 1995, pages 483-512, XP001073411	
A	WO 91 00905 A (US HEALTH) 24 January 1991 (1991-01-24)	
A	WO 00 09734 A (INTRONN HOLDINGS LLC) 24 February 2000 (2000-02-24)	
A	WO 97 22250 A (INTRONN LLC ;MITCHELL LLOYD G (US)) 26 June 1997 (1997-06-26) figure 2	
P, X	WO 01 32896 A (ALEXANDROV KIRILL ;JENA BIOSCIENCE GMBH (DE); GRUN MATHIAS (US)) 10 May 2001 (2001-05-10) figure 3	1-5, 8-10, 13

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PC1/EP 01/05794

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0001846 A	13-01-2000	AU 4907999 A	24-01-2000
		BR 9911802 A	22-01-2002
		CN 1323354 T	21-11-2001
		DE 1093526 T1	11-10-2001
		WO 0001846 A2	13-01-2000
		EP 1093526 A2	25-04-2001
		GB 2349885 A	15-11-2000
		GB 2362885 A	05-12-2001
		NO 20010019 A	05-03-2001
WO 9937770 A	29-07-1999	AU 2240899 A	09-08-1999
		AU 4608499 A	30-12-1999
		BR 9911097 A	13-02-2001
		CA 2318935 A1	29-07-1999
		CN 1305525 T	25-07-2001
		WO 9964586 A2	16-12-1999
		EP 1049776 A1	08-11-2000
		EP 1084242 A2	21-03-2001
		HU 0101593 A2	28-08-2001
		HU 0102349 A2	28-10-2001
		PL 342082 A1	21-05-2001
		PL 345951 A1	14-01-2002
		WO 9937770 A1	29-07-1999
WO 9100905 A	24-01-1991	US 5135855 A	04-08-1992
		AT 157701 T	15-09-1997
		AU 642657 B2	28-10-1993
		AU 6034490 A	06-02-1991
		CA 2063427 A1	08-01-1991
		DE 69031389 D1	09-10-1997
		DE 69031389 T2	19-03-1998
		EP 0484374 A1	13-05-1992
		ES 2109924 T3	01-02-1998
		JP 2525289 B2	14-08-1996
		JP 5501048 T	04-03-1993
		WO 9100905 A1	24-01-1991
		US 5126251 A	30-06-1992
WO 0009734 A	24-02-2000	US 6083702 A	04-07-2000
		US 6280978 B1	28-08-2001
		AU 5673599 A	06-03-2000
		EP 1105513 A2	13-06-2001
		WO 0009734 A2	24-02-2000
WO 9722250 A	26-06-1997	US 6013487 A	11-01-2000
		AU 1329997 A	14-07-1997
		WO 9722250 A1	26-06-1997
		US 6083702 A	04-07-2000
		US 6280978 B1	28-08-2001
		AU 730305 B2	01-03-2001
		CA 2240494 A1	26-06-1997
		EP 0883344 A1	16-12-1998
		JP 2000515362 T	21-11-2000
WO 0132896 A	10-05-2001	AU 1997901 A	14-05-2001
		WO 0132896 A1	10-05-2001

Form PCT/ISA/210 (patent family annex) (July 1992)

NUC 38929

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.